Tema

4

LA PERCEPCIÓN DEL LENGUAJE

1.- INTRODUCCIÓN.

El lenguaje consta de 2 clases de fenómenos:

- **unos públicos**, y como tal observables y medibles, y
- **otros privados**, y por ello inferidos y supuestos por el investigador.

En su dimensión pública y observable, el lenguaje natural es un estímulo físico compuesto por sonidos percibidos auditivamente o trazos percibidos visualmente.

En su dimensión privada, el lenguaje natural es una representación mental dotada de significado.

Para que la comunicación lingüística entre dos interlocutores sea posible, ***es preciso que el hablante/escritor codifique sus significados en sonidos de habla (o trazos de escritura) y que el oyente/lector descodifique estos sonidos o trazos en significados.***

Por consiguiente, la primera tarea que un oyente o un lector tiene que llevar a cabo en la comunicación verbal ***es la de transformar los estímulos físicos de habla o escritura que recibe en una representación de los sonidos o las letras que componen la emisión verbal.***

La percepción del lenguaje oral o escrito **es un proceso que se realiza sin apenas esfuerzo consciente por parte del perceptor y con notable eficacia en la mayoría de las ocasiones,** incluso en condiciones en la señal física nos llega distorsionada, como cuando percibimos habla susurrada o a través de un teléfono o un altavoz (o en la escritura manual), o enmascarada por ruido, como sucede en muchos de los intercambios comunicativos habituales.

Por otra parte**, la percepción del lenguaje se mantiene relativamente constante aun cuando ciertas propiedades físicas de la estimulación varíen**; así, por ejemplo, reconoceremos como idénticas 2 secuencias de habla a pesar de las variaciones que se puedan producir en la frecuencia de la voz del hablante (vg., voz femenina o masculina), en la velocidad del habla o a través de diferentes acentos regionales.

Estas observaciones iniciales nos dan a entender que **el sistema humano de percepción del lenguaje tiene que ser un sistema altamente flexible y adaptativo, a la vez que considerablemente automático.**

Por otra parte, hay que subrayar que la percepción del lenguaje oral y del lenguaje escrito presentan diferencias nada desdeñables:

- **La más obvia es que cada una de ellas afecta a un órgano sensorial distinto: la audición y la visión, respectivamente**.

- Pero, además**, la percepción del habla tiene que hacer frente a unas limitaciones temporales de las que se ve libre la percepción de la escritura**. Dado que

  - **el estímulo auditivo del habla** se desvanece de forma inmediata, no podemos tener un acceso permanente a él, por lo que hemos de registrarlo y procesarlo a la mayor brevedad posible; en cambio,

  - **el estímulo visual del lenguaje escrito** permanece a nuestro alcance todo el tiempo que sea necesario, por lo que no impone una exigencia tan apremiante de procesamiento inmediato. Esto permite al lector efectuar retrocesos y correcciones en presencia del estímulo.

- Una tercera diferencia entre ambas modalidades sensoriales del lenguaje es **que el habla es un estímulo continuo**, con pocos cortes o silencios entre unidades lingüísticas que el oyente tiene que descodificar, es decir, cada sonido, sílaba o palabra casi nunca va separado por pausas, mientras que **la escritura se nos ofrece dividida en unidades**, palabras, en el caso de la escritura manual, y palabras y letras, en el de la escritura impresa.

El objetivo primordial de este capítulo es examinar cómo se realizan los procesos de percepción o reconocimiento del lenguaje humano, es decir, cuáles son los mecanismos sensoriales, cognitivos y lingüísticos implicados en el procesamiento de las señales físicas del lenguaje.

Para ello, nos centraremos principalmente en los procesos de percepción del lenguaje oral, haciendo especial hincapié en las primeras etapas de estos procesos, es decir, en la identificación de las unidades lingüísticas más elementales (sonidos o fonemas individuales) que forman las palabras de la lengua.

Debemos tener en cuenta, no obstante, **que el lenguaje no se percibe en forma de unidades fragmentadas, sino como una sucesión continua de elementos perceptivos**, es decir, si bien parece cierto que las etapas iniciales de la percepción del lenguaje constan de procesos casi exclusivamente dirigidos por las propiedades de la señal (de "abajo-arriba"), también hay que tomar en consideración otros factores contextuales (vg., representaciones léxicas, estructuras sintácticas y conocimiento semántico) que guían, a partir de un determinado momento, estas operaciones perceptivas.

Por otra parte, a la vista de las diferencias que hay entre las condiciones de percepción del lenguaje en las modalidades auditiva y visual, dedicaremos un último apartado a examinar los procesos de percepción del lenguaje escrito y los mecanismos empleados en la extracción de la información visual durante la lectura.

2.- PRINCIPALES PROBLEMAS DE LA PERCEPCIÓN DEL HABLA.-

La percepción del habla se puede definir como el proceso en virtud del cual un patrón de variación de energía acústica, que incide en los órganos receptores de la audición, se transforma en una representación mental de la configuración estimular (i.e., los fonemas o sonidos) que produce esa variación de energía.

Desde el punto de vista físico,

- el habla, lo mismo que cualquier otro estímulo auditivo, **se define como onda sonora con propiedades acústicas** (vg., frecuencia y amplitud) **y temporales** (vg., duración).

- Estas propiedades de la onda sonora son, en último término, producto de cambios o variaciones en la presión que se difunden en el espacio. La tarea de nuestro sistema perceptivo es, por consiguiente, **extraer de estas variaciones que se producen en la onda sonora unas constancias perceptivas que, en el caso del habla, corresponden a las unidades lingüísticas que llamamos fonemas.**

- Dado que estas unidades **se perciben ordenadas en una secuencia temporal**, es de suponer que también **las claves acústicas presentes en la señal de habla han de estar temporalmente ordenadas.**

Así pues,

la tarea de percibir habla podría definirse, a primera vista, como la sucesiva transformación, de uno en uno, de segmentos de la onda sonora del habla en fonemas individuales. Se trataría, por consiguiente, de una tarea de reconocimiento de patrones, como otras tantas de las estudiadas en psicología cognitiva, en la que un conjunto de rasgos físicos se haría corresponder con una representación mental almacenada en la memoria del oyente. Nada hay, sin embargo, más lejos de la realidad.

La percepción del habla es una actividad bastante compleja y con intrigantes paradojas. En este apartado vamos a exponer los fenómenos responsables de esta complejidad, fenómenos que, paradójicamente, hacen posible al mismo tiempo que este complejo proceso pueda llevarse a cabo con notable eficacia.

Para entender estos fenómenos es necesario examinar antes las propiedades físicas de los sonidos del habla.

☒ Mediante el procedimiento conocido como **espectrografía de sonidos** es posible obtener una representación visual del habla, o espectrograma, que recoge la composición de frecuencias de la voz en unidades de tiempo.

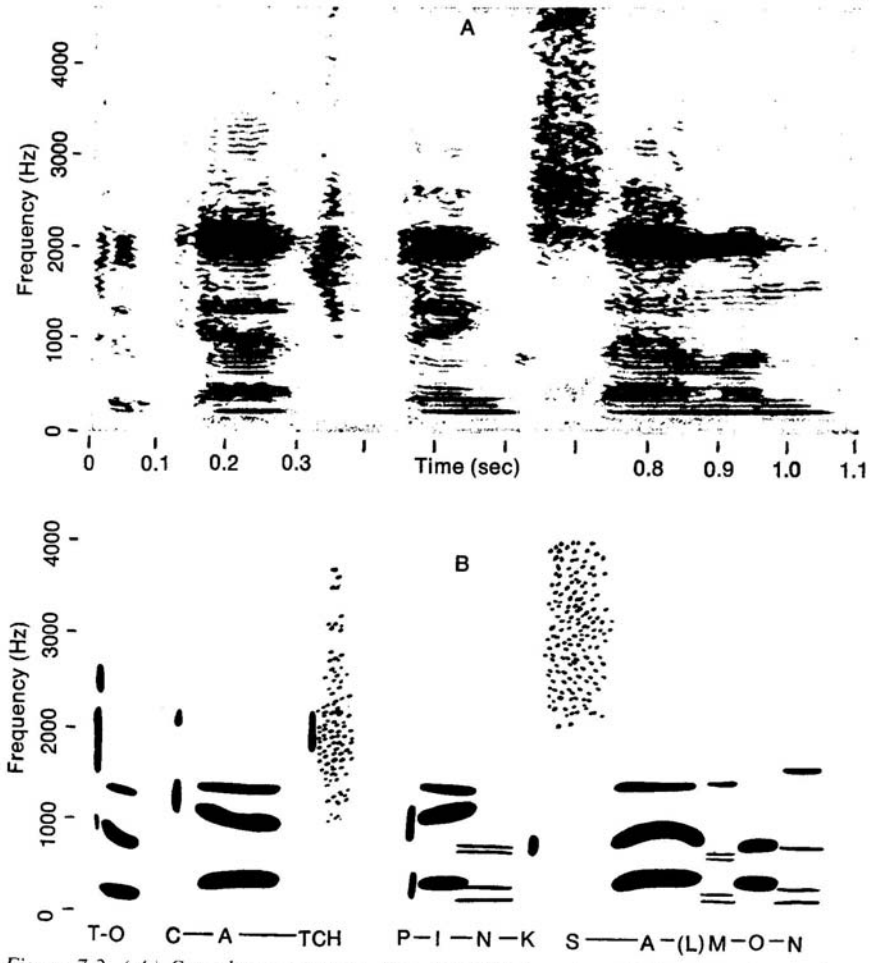


Figure 7.2. (A) Sound spectrogram. (B) Simplified version of the same phrase that serves as input to pattern playback. (After Cooper et al., *1952.)*

☒ Así, como muestra la figura 8.1, **algunas secuencias de sonidos de habla se descomponen en unas bandas de frecuencia** (expresada en "hercios" (Hz)) denominadas **<u>formantes</u>** (representadas en el eje Y o de ordenadas), que se numeran a partir de los valores más bajos de frecuencia: en nuestro ejemplo,

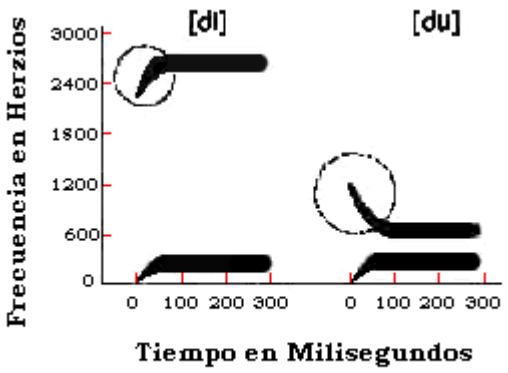


Figura 8.1

- el primer formante de ambas sílabas arranca desde los 0 Hz y asciende hasta situarse en torno a los 300 Hz;
- el segundo formante de la sílaba /di/ (izquierda) asciende de 2.400 a 2.600 Hz, mientras que el de la sílaba /du/ (derecha) desciende de 1.200 a 600 Hz.
- Por otra parte, **el tiempo medido**, en este caso, en segundos, se representa en el eje-X o de abcisas.
- Cada formante consta de dos partes:

  - una parte curva, llamada "transición del formante", en la que el valor de la frecuencia va modificándose progresivamente, y
  - una parte recta o "parte estable del formante".

La figura muestra los espectrogramas de las sílabas /di/ (izquierda) y /du/ (derecha)*, pronunciadas en inglés.

En líneas generales, podría pensarse que la transición de los formantes corresponde a la consonante de la sílaba y la parte estable de los formantes a la vocal.

(*) Esta es una representación artificial y, por tanto, simplificada, del espectograma de dichas sílabas. En ella se han conservado únicamente los formantes primero y segundo, que son suficientes para discriminar una sílaba de la otra. Los espectrogramas "normales" presentan bandas de frecuencia considerablemente más gruesas y difuminadas.

Por consiguiente, se diría que existe una correspondencia entre las propiedades físicas de cada sonido y el resultado perceptivo.

Sin embargo, a tenor de lo reflejado en las gráficas, es evidente que este no es el caso, dado que los valores de transición del segundo formante difieren notablemente de una sílaba a otra, aun cuando la consonante inicial /d/ sea la misma en ambos casos.

En otras palabras, hay una diferencia sustancial entre las claves acústicas que nos

sirven para identificar el mismo fonema /d/ en dos contextos fonéticos dispares.

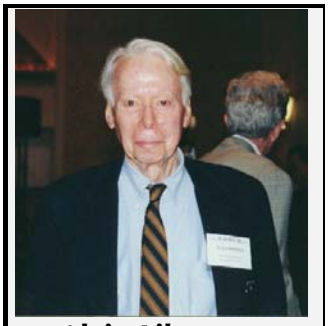
Esto indica **que la percepción no es un proceso de "traducción directa" de propiedades o claves acústicas a representaciones fonémicas**, es decir, **cada fonema de la lengua no corresponde siempre al mismo conjunto de unidades acústicas**.

Esta falta de correspondencia acústico-fonémica se manifiesta en dos problemas distintos, aunque estrechamente relacionados entre sí:

- **la segmentación y**
- **la ausencia de invarianza.**

El problema de la segmentación

- **Está relacionado con el hecho de que la señal del habla es continua, mientras que los sonidos del habla se perciben como discretos o discontinuos.**
- Esta percepción discreta se da a pesar de que las claves acústicas necesarias para identificar fonemas no se distribuyen de manera separada a lo largo de fonemas adyacentes (/d/ e /i/, o/d/ y /u/, en el ejemplo anterior), sino que aparecen "mezcladas" en sonidos próximos entre sí.
- Prueba de ello es la observación de que al separar de manera artificial el fragmento del espectro que corresponde a la consonante de una sílaba del que corresponde a la vocal de la misma sílaba y al presentar únicamente la parte de la consonante, los oyentes eran capaces de "adivinar" qué vocal seguía a dicha consonante. Esto obedece a que cada fragmento de habla retiene información acústica del fragmento contiguo. **Dicho de otro modo, las claves acústicas transmiten información en paralelo sobre segmentos fonéticos sucesivos** (cfr. Liberman et al., 1967).



Alvin Liberman

El problema de la ausencia de invarianza,

- Relacionado con el anterior, **se define como la falta de correspondencia biunívoca entre fragmentos de la señal acústica y fonemas discretos**.
- Los segmentos de habla carecen, en su mayoría, de propiedades invariantes, ya que, a causa de la naturaleza continua del estimulo de habla, se ven influidos por el contexto acústico en que se encuentran.
- Pese a ello, los oyentes somos capaces de descubrir constancias perceptivas y con ello identificar eficazmente sonidos de habla. Por ejemplo, nuestro sistema perceptivo identifica el mismo fonema (/n/) aun cuando éste se encuentre en contextos acústicos sustancialmente diferentes, como sucede en las palabras "co[n]trario", "con[n]ato" o "co[n]vocatoria" (cfr., Pisoni y Luce, 1987; Miller, 1990).

Los problemas de segmentación y de ausencia de invarianza

- Tienen un mismo origen, a saber, **las demandas "co-articulatorias" del habla.**
- Cuando pronunciamos una secuencia de fonemas, no articulamos cada fonema por separado, sino que ajustamos la posición de nuestros órganos articulatorios a las configuraciones que éstos han de adoptar para producir los fonemas inmediatamente anterior y posterior.
- Así, al articular el fonema /n/ de los ejemplos anteriores,
  - la posición de la lengua es más extendida y plana en "co[n]trario",
  - algo más retraída y en contacto con los alvéolos en "co[n]ato",
  - y se halla en reposo en "co[n]vocatoria" (hasta el punto de que el fonema que articulamos es más una consonante bilabial [m] que una alveolar [n]).
- Estas diferencias vienen determinadas, en estos casos, por el contexto articulatorio subsiguiente (una consonante dental /t/, en "co[n]trario", una vocal /a/, en "co[n]ato", y una consonante bilabial /b/, en "co[n]vocatoria").

En suma, los movimientos articulatorios efectuados en fonemas sucesivos se solapan en el tiempo, lo que origina diferencias acústicas en un mismo fonema en función del contexto acústico-articulatorio en que éste se localiza.

Estas demandas coarticulatorias son también responsables de que el habla humana se emita a considerable velocidad, hasta el punto de que la tasa normal de emisión de fonemas por unidad de tiempo **(que oscila entre 10 y 15 fonemas por segundo, aunque pueda alcanzar de 25 a 30 en condiciones de habla muy rápida)** superaría el poder de resolución del oído humano si cada fonema estuviese realmente asociado a un conjunto de claves discretas y específicas.

Algunos intentos de traducir signos gráficos a un alfabeto acústico para proporcionar un mecanismo de lectura sonora para sujetos ciegos **(*)** (cfr. Liberman et al., 1967; Liberman, 1982) han terminado en fracaso, debido precisamente a las limitaciones perceptivas de nuestro sistema auditivo.

(*) En su artículo de 1982, Liberman relata las experiencias fallidas llevadas a cabo en los laboratorios Haskins en los años 50 y 60 para diseñar una máquina de lectura para ciegos. Este ingenio hubiera consistido (pues nunca llegó a construirse) en un mecanismo capaz de inspeccionar letras impresas y traducir cada letra a un patrón sonoro distinto de los de otras letras. El usuario tendría que aprender a asociar cada sonido con cada letra y, con el tiempo, llegaría a "leer" las secuencias de sonidos producidas por la máquina de manera fluida.

Esto viene a apoyar, una vez más, **la idea de que el sistema perceptivo para el habla no puede estar basado en correspondencias biunívocas entre claves acústicas y unidades fonémicas.**

Los problemas mencionados hasta aquí no agotan ni mucho menos las dificultades a que se enfrenta el sistema perceptivo en la descodificación del habla humana. Entre los problemas que no vamos a analizar aquí, **cabe resaltar la presencia indisociable en toda emisión verbal de características acústicas relevantes e irrelevantes**; entre estas últimas figuran, por ejemplo,

- La frecuencia fundamental general de la voz, que establece lo que vulgarmente llamamos el "tono" de voz del hablante (vg., voz femenina o masculina), o
- La intensidad del habla.

Asimismo, **la percepción del habla se mantiene, en gran medida, intacta aun cuando se produzca una pérdida de información acústica por efecto del ruido.**

¿Cómo se resuelve nuestro sistema perceptivo tan arduos problemas de un modo tan directo y eficaz?

En los dos apartados que siguen se va a exponer algunas de las respuestas que se han ofrecido a este interrogante.

3.- PROCESOS BÁSICOS DE LA PERCEPCIÓN DEL HABLA.-

En este apartado se va a examinar algunos de los **fenómenos más relevantes de la percepción del habla en sus primeras etapas.**

Para ello se comenzará por describir sumariamente estas etapas para centrarse, a continuación, en la propiedad tal vez más llamativa de la percepción del habla en sujetos humanos, tanto bebés como adultos, a saber**, su naturaleza categorial**.

Esto servirá, por su parte, para plantear hasta qué punto **la percepción del habla se basa en procesos psicológicos** especialmente diseñados para realizar esta tarea o, por el contrario, descansa en **mecanismos auditivos inespecíficos**, empleados en la percepción de cualquier clase de estímulos acústicos.

3.1.- Etapas en la percepción del habla.

Las etapas en que se divide el proceso de percepción del habla son, en principio**, de naturaleza lógica o teórica**, es decir, se apoyan en consideraciones **más lingüísticas que psicológicas.**

Por esta razón, no se postularon inicialmente como estadios temporalmente discretos y funcionalmente autónomos, sino más bien como una caracterización de las sucesivas transformaciones que sufre la información desde que, como señal física de habla, llega a los órganos receptores hasta que, en un formato más abstracto (i.e. fonológico) se emplea en procesos superiores de acceso al léxico y comprensión del lenguaje.

No obstante, a pesar de su carácter teórico, hay evidencia empírica a favor de la división de los procesos de percepción del habla en **4 etapas**. Estas son (cfr. Studdert-Kennedy, 1974, citado en Pisoni y Luce, 1987; Clark y Clark, 1977):

- **Análisis auditivo periférico.**
- **Análisis auditivo central.**
- **Análisis acústico-fonético y,**
- **Análisis fonológico.**

1.- Análisis auditivo periférico:

En esta primera etapa, **se produce una descodificación preliminar de las señales de habla en el sistema auditivo periférico** o, lo que es lo mismo, las estructuras anatómicas y neurales del oído realizan un primer análisis de las propiedades acústicas relevantes de la señal.

Los mecanismos de descodificación investigados en este nivel son de dos clases:

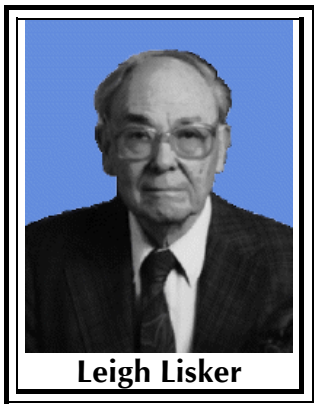
- por una parte, **mecanismos neuroacústicos**, como por ejemplo patrones de descarga de fibras nerviosas que sintonizan con atributos de la señal de habla que se hallan presentes, por ejemplo, en las consonantes oclusivas de inicio de sílaba (vg., /pa/, /ba/, /ta/, /da/, /ga/, /ka/;
- y por otra, **mecanismos psicoacústicos** definidos en términos más

abstractos, esto es, con independencia de sus correlatos fisiológicos. Un ejemplo de estos mecanismos son los llamados *"filtros paso-banda"*, que efectúan transformaciones de la señal mediante el análisis de algunos de sus componentes (vg. La resolución de las frecuencias más bajas del espectro, para separar el 1º y el 2º formante, y de las más elevadas, para determinar el valor temporal de las explosiones, en las consonantes oclusivas).

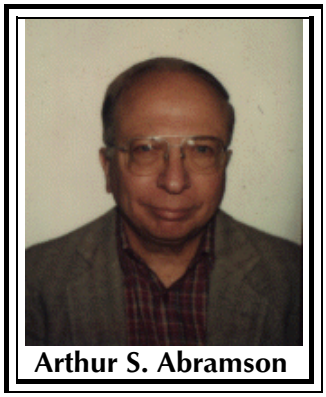
En estudios relativamente recientes se ha podido constatar que el sistema auditivo periférico humano posee **unas capacidades de filtrado de las señales acústicas**, que, aunque no sirven por sí solas para identificar fonemas, contribuyen a aislar algunos componentes básicos de la señal de habla y con ello a reducir el problema de la ausencia de invarianza anteriormente aludido.

2.- Análisis auditivo central:

El cometido de este análisis **es extraer de la señal una serie de patrones espectrales** (vg., frecuencia fundamental, dirección de las transiciones de los formantes) **y temporales** (vg., el desfase entre ciertos eventos que ocurren durante la emisión de la señal) **y almacenarlos en una memoria auditiva de muy breve duración** (la llamada "memoria ecoica"). Del análisis de estos patrones se obtienen unas "claves acústicas" o propiedades que se combinan para dar lugar a los fonemas.


Leigh Lisker

No obstante, como ya se ha señalado, la relación entre las claves acústicas y los segmentos fonéticos no siempre es directa, esto es, cada segmento fonético no siempre está constituido por las mismas claves acústicas. Así, por ejemplo, la sonoridad de las consonantes oclusivas iniciales de sílaba (vg., que permite distinguir entre sonoras como /ba/, /da/, /ga/, y sordas como /pa/, /ta/, /ka/) viene dada indistintamente por claves como:


Arthur S. Abramson

- el "tiempo de emisión de voz" (véase punto 3.2),
- la "frecuencia de la transición del primer formante" o
- la "amplitud de la explosión –Burst – (Abramson y Lisker, 1965; Suárez Buratti, 1991; para una revisión, véase (Pisoni y Luce, 1987).

En suma, **las claves acústicas que nos permiten identificar propiedades fonéticas son, en su mayoría dependientes del contexto acústico**.

No obstante **algunas claves acústicas sí parecen estar directamente asociadas con rasgos fonéticos específicos**; se trata, en este caso, de claves independientes de contexto o propiedades invariantes de la señal (Blumstein y Stevens, 1979, 1980).

Entre ellas cabe destacar **el "ruido fricativo"** presente en consonantes "estridentes" (vg., /s/, /z/, /š/, /ž/, /č/), o ciertas configuraciones espectrales asociadas al punto de articulación de las consonantes (vg., relevantes para la distinción entre labiales /b/, alveolares /d/ y velares /g/).

3.- Análisis acústico-fonético:

- **Este es el primer nivel de análisis en que se efectúa un procesamiento propiamente lingüístico de la señal.**
- El objetivo de esta etapa **es identificar los segmentos o fonemas del habla**. En ella las claves acústicas se acoplan a los rasgos distintivos fonéticos, esto es, a las propiedades acústico-articulatorias en que se descomponen los fonemas de la lengua.

Los rasgos fonéticos son, por tanto, representaciones abstractas que sirven de mediadores entre los planos físico (acústico) y lingüístico (fonético).

- Asimismo, en esta etapa **se verifica la categorización perceptiva del habla**, es decir, se descubren las constancias perceptivas que nos permiten identificar sonidos discretos, resolviéndose los problemas de segmentación y variabilidad anteriormente aludidos.

Algunos investigadores (vg., Miller y Nicely, 1955; Eimas y Corbit, 1973) **han propuesto la existencia en este nivel de unos "detectores de rasgos",** o mecanismos neurales especializados en la identificación de rasgos fonémicos distintivos (vg., sonoridad, nasalidad, estridencia, duración, etc.) (Clark y Clark, 1977, para una descripción de estos detectores de rasgos).

En el punto 3.4 de este capítulo se examinará la evidencia aducida por algunos investigadores a favor de la existencia de estos detectores.

4.- Análisis fonológico:

- En esta última etapa, **los rasgos y segmentos fonéticos identificados en la etapa anterior son convertidos en segmentos fonológicos**, es decir, en representaciones abstractas de los sonidos que se someten a reglas combinatorias para formar unidades superiores, tales como las sílabas y las palabras.

- En este nivel, **ciertas distinciones fonéticas** (vg., la distinción entre la oclusiva /b/ inicial de palabra, vg., #/[b]/esar#, o interna a una palabra, vg., #sa/[b]/er#) **dejan de ser fonológicamente relevantes para convertirse en variaciones alofónicas del mismo fonema** (i.e., el mismo sonido con dos variantes o "alófonos".

- Asimismo, ciertos **fenómenos de asimilación o transformación fonética que aparecen vinculados a procesos de derivación morfológica** (vg.,

  - ♣ el cambio $/\Theta/ \rightarrow /k/$ en $\#electri/\Theta/idad\# \rightarrow \#eléctri/k/o\#$, o

  - ♣ $/x/ \rightarrow /g/$ en $\#ma/\times/ia\# \rightarrow ma/g/o\#$)

  se explican en virtud de reglas fonológicas que operan en este nivel.

El resultado del análisis fonológico es una secuencia lineal de fonemas organizados en una estructura jerárquica de constituyentes fonológicos.

Esta organización jerárquica viene dada por la estructura de la sílaba, que comprende los siguientes elementos constituyentes:

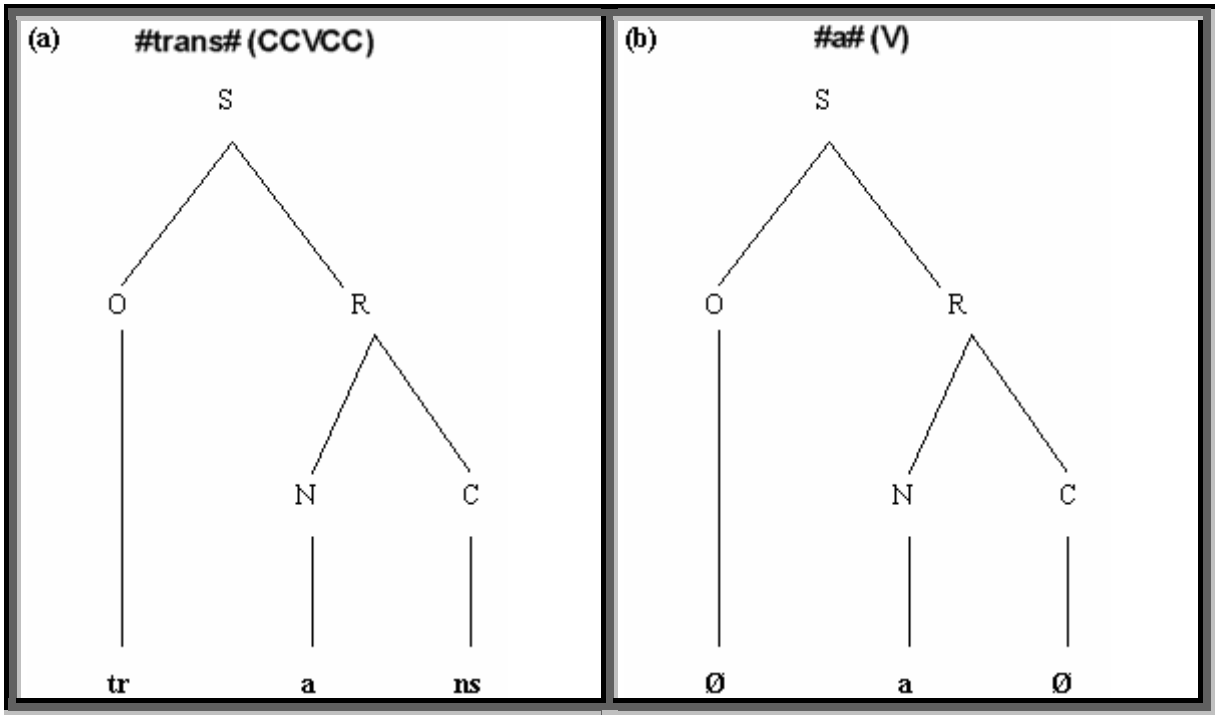
1) **el "inicio" u onset,** formado por una consonante o grupo consonántico inicial de carácter opcional, y
2) **la "rima" o rime,** que a su vez se divide en dos constituyentes menores,

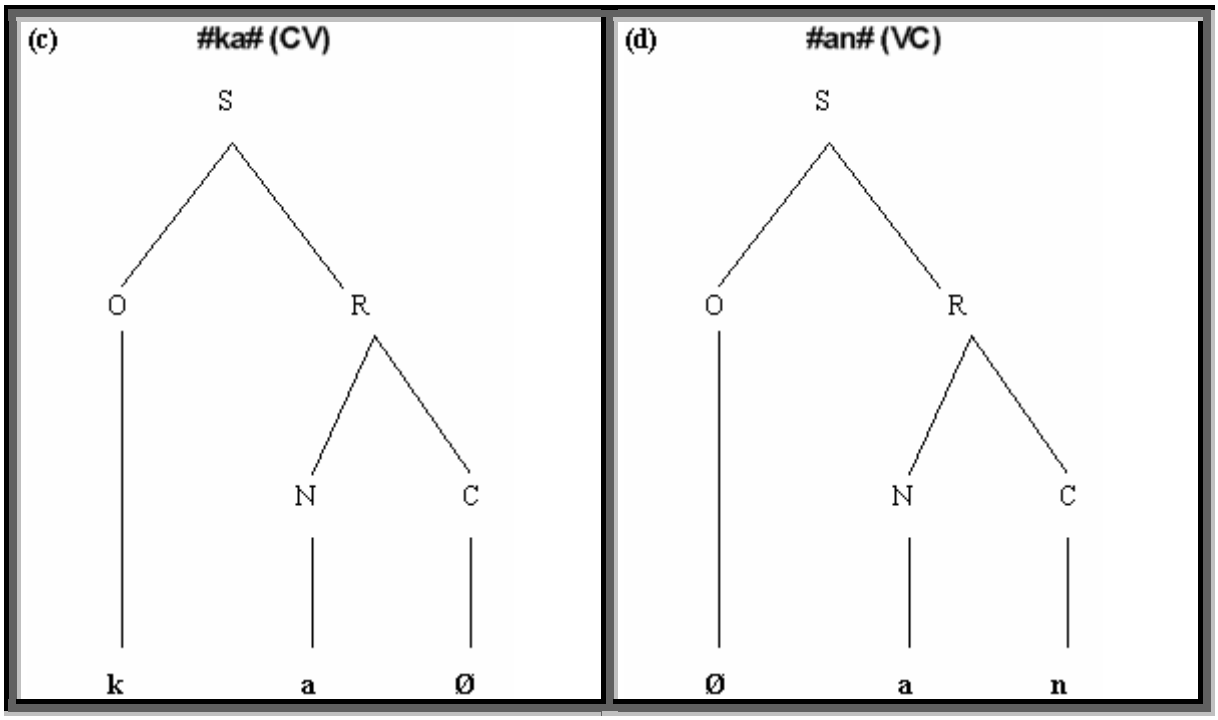
   - ♣ ***el "núcleo" vocálico*** *y*
   - ♣ ***la "coda" o terminación consonántica***, también opcional (cfr. Treiman, 1983)

La figura 8.2 muestra la estructura de algunas sílabas del castellano:

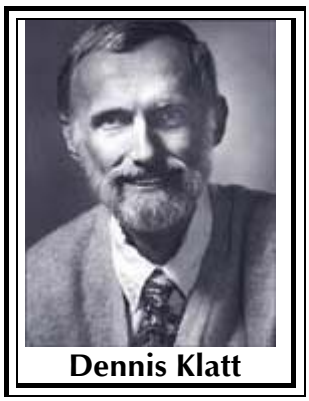
a) #trans#, en "transporte";
b) #a#, en "ave";
c) #ka#, en "casa"; y
d) #an#, en "antiguo".

Figura 8.2.- Algunos ejemplos de estructuras silábicas del castellano, correspondientes a las sílabas #trans# (CCVCC), #a# (V), #ka# (CV) y #an# (VC). (S = sílaba; O = Onset; R = rima; N = núcleo; C = coda).

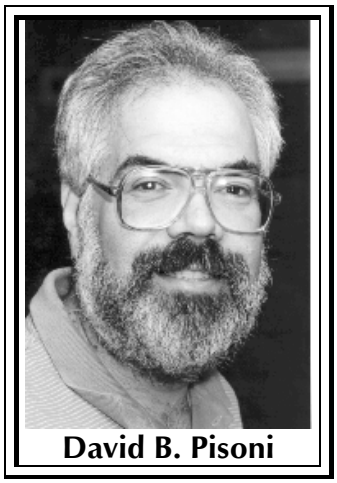




Como señalábamos al comienzo de este punto, **no parece haber acuerdo la realidad psicológica de cada uno de estos niveles de procesamiento en la percepción del habla,** como tampoco sobre el curso temporal de estos 4 procesos y sus posibles interacciones.


Dennis Klatt

Por ejemplo, hay autores **(Klatt, 1980)** que **consideran inadecuado postular un nivel independiente de representación fonética**, debido a la dificultad de hallar un acoplamiento sistemático entre claves acústicas y segmentos fonéticos. Por ello, optan por aplazar la resolución del problema de la ausencia de invarianza hasta un nivel superior de procesamiento, es decir, hasta los procesos de acceso al léxico (cfr. Capítulo 9).


David B. Pisoni

Como han observado **Pisoni y Luce (1987**), **si se elimina el nivel fonético, tampoco habrá razones para postular un nivel fonológico de procesamiento, dado que en tal caso careceríamos de la representación de entrada a este nivel**.

Quienes así piensan consideran que las representaciones fonética y fonológica son constructos lingüísticos que no juegan un papel psicológicamente necesario durante el procesamiento de la señal, sino en todo caso a posteriori, es decir, sólo si y cuando el oyente hace uso de este tipo de información para efectuar determinadas tareas (vg., detección de fonemas en contextos experimentales, uso de información fonética en tareas de discriminación perceptiva, lectura de palabras desconocidas o poco frecuentes, etc.).

Pese a todo, **vamos a ignorar por el momento este punto de vista y a centrarnos en las explicaciones de la percepción del habla que sí postulan procesos de transformación acústico-fonética.**

Es preciso recordar, en este punto, que las así denominadas **"representaciones fonémicas"** no se encuentran presentes en la señal de habla, sino que, de algún modo, son suministradas por el propio perceptor a partir de información de que dispone en su memoria.

Así, pues, aunque los procesos de transformación acústico-fonética estén dirigidos inicialmente por propiedades de la señal (es decir, aun cuando sean procesos de "abajo-arriba"), también dependen del uso de información "superior", es decir, transmitida de "arriba-abajo".

Entre las pruebas comúnmente aducidas a favor de la existencia de procesos de transformación acústico-fonética, cabe resaltar las investigaciones sobre la "percepción categorial" de segmentos fonéticos y sobre los mecanismos de detección de rasgos.

3.2.- La percepción categorial

En las tareas perceptivas habitualmente utilizadas en psicología, al sujeto se le puede pedir que efectúe dos clases de juicios:

- **juicios de discriminación**, consistentes en señalar la diferencia entre dos o más estímulos x e y;
- **juicios de identificación**, que consisten en indicar si un estímulo x pertenece o no a una categoría X.

En ambos casos, lo que se le pide al sujeto es que efectúe un **juicio categorial**, es decir, *que clasifique un estímulo dentro de una determinada categoría*. La razón por la que se emplean estas tareas y no otras (***3****) es que las representaciones almacenadas en su memoria con las que el sujeto compara los estímulos que recibe son categorías separadas o discretas, es decir, no existe un continuo que abarque varias o todas las diversas categorías fonémicas de la lengua.

Sin embargo, desde hace tiempo se sabe que ciertas dimensiones físicas de los sonidos del habla (i.e., ciertas claves acústicas) **se definen como variables continuas.**

Así, la diferencia entere los sonidos pertenecientes a dos categorías fonémicas distintas (vg., consonantes sonoras, como /b/, y sordas, como /p/) es una cuestión de grado o, lo que es lo mismo, existe una dimensión continua de variación en uno de cuyos extremos se define una categoría fonémica y en el otro otra.

En el caso de la distinción entre consonantes sonoras y sordas, esta dimensión de variación es el llamado **"tiempo de emisión de voz" (TEV, o voice onset time, VOT, en inglés),** un parámetro que se define *como el desfase temporal existente entre dos sucesos articulatorios que son por una parte, la liberación de la oclusión o constricción del tracto vocal producida al articular la consonante, y por otra, la emisión de la voz o "ataque glótico", o sea, la vibración de las cuerdas vocales*.

El TEV es una dimensión continua, dado que oscila gradualmente entre un valor negativo (vg., 150 mseg.) y otro positivo (vg., +150 mseg.). Sin embargo nuestra percepción de la diferencia entre consonantes sonoras (con un TEV negativo o positivo pero bajo) y sordas (con un TEV positivo y elevado) no es continua y gradual, sino discreta y abrupta, según se aprecia en las curvas de discriminación (izquierda) e identificación (derecha) de la figura 8.3.

En una tarea clásica de discriminación, se le presenta al sujeto un par de estímulos auditivos (dos sílabas CV) y se le pide que indique si son iguales o diferentes, esto es, un juicio de discriminación perceptiva; la tarea de identificación, por su parte, consiste en "nombrar" el estímulo que el sujeto ha escuchado, en este caso también una sílaba CV.

(3*) La tarea de identificación consiste en una tarea de etiquetado de estímulos esto es, en indicar a cuál de 2 categorías previamente establecidas corresponde el estímulo presentado.

La tarea de discriminación, también denominada tarea ABX, consiste en presentar al sujeto secuencialmente un par de estímulos A y B y a continuación un tercer estímulo X, pidiendo al sujeto que señale a cuál de los dos estímulos previos era igual el tercero.

En estudios más recientes se han introducido tareas algo más minuciosas, que no sólo incluyen juicios categoriales de discriminación e identificación, sino también juicios subjetivos de "confianza", en los que se pide al sujeto que señale el grado de seguridad con que clasifica cada estímulo.

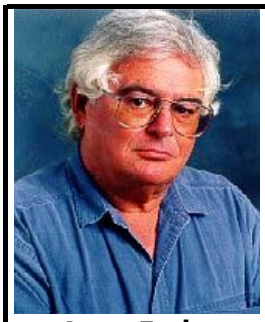
Por otra parte, se han diseñado tareas considerablemente ingeniosas para someter a prueba las capacidades perceptivas de organismos desprovistos de lenguaje, como por ejemplo bebés humanos o animales de otras especies, con estímulos de habla humana.

En los hablantes del inglés, el valor crítico de TEV para discriminar consonantes sonoras y sordas se sitúa en torno a los +25 mseg.; en los del castellano, en cambio, adopta un valor algo más bajo (entre +5 y +10 mseg.) (cfr. López Bascuas, 1991).

A la vista de estas observaciones, cabe formular algunas preguntas. En primer lugar, una pregunta de carácter general: **¿por qué categorizamos los sonidos del habla de esta manera?, ¿de qué depende la percepción categorial del habla?** Esta pregunta da paso a otras más específicas, aunque no menos importantes:

1) **¿En qué medida esta forma de categorizar los sonidos del habla depende de nuestra experiencia con el lenguaje, ¿se trata de una habilidad aprendida o de una capacidad innata?;**
2) **¿Hasta qué punto la percepción categorial se aplica exclusivamente a estímulos de habla o, por el contrario, caracteriza también nuestra percepción de los continuos de no habla (vg., estímulos musicales, zumbidos, etc.)?;**
3) **¿En qué medida la percepción categorial del habla es privativa de la especie humana o compartida por otras especies animales?**

Estas 3 preguntas nos sitúan ante la posibilidad, actualmente en debate, de que los procesos de percepción del habla formen un conjunto de mecanismos especialmente diseñados para percibir el habla humana, completamente independientes de los restantes mecanismos al servicio de la percepción auditiva.


Jerry Fodor

En otras palabras, puede postularse que **la percepción del habla es una facultad vertical al estilo de las propuestas por la teoría de la modularidad de Fodor** (1983), es decir, definidas en virtud del tipo del tipo de información con la que operan. Dentro de este marco general de discusión

- **la primera** pregunta se refiere a la cuestión del innatismo "computacional" de la percepción del habla,
- **la segunda,** a la especificidad de dominio del sistema de percepción del habla, y
- **la tercera,** a la especificidad humana de este sistema perceptivo.

En primer término, vamos a revisar algunos datos relevantes para las 3 preguntas específicas, para luego centrarnos en la pregunta general que formulábamos al inicio de este párrafo.

3.3. La percepción del habla en bebés.

Las investigaciones sobre la percepción del habla en bebés **se han centrado fundamentalmente en el estudio de habilidades de categorización de continuos de habla,** con el doble propósito de averiguar:

- Hasta qué punto estas habilidades se hallan presentes desde edad muy temprana y
- Si los bebés comparten con los sujetos adultos los mismos valores críticos de discriminación.

Los estudios más conocidos al respecto son los llevados a cabo por Peter Eimas y colaboradores en los años 70 (Cutting y Eimas, 1975; Eimas, 1985).

- Estos investigadores deseaban comprobar si bebés de pocas semanas de edad (de 1 y 4 meses) dividían continuos de habla (vg., /ba/ a /pa/) en categorías discretas y si lo hacían en torno al valor de TEV de -20, 0, +20, +40, +60 y +80 mseg.

- La tarea empelada en los experimentos (que aparece descrita con detalle en Eimas, 1985), sumamente ingeniosa por lo demás, era **una tarea de discriminación** que se registraba de forma indirecta, dada la incapacidad del bebé para manifestar verbalmente sus juicios perceptivos.

- El procedimiento empleado, llamado **"succión de elevada amplitud",** consistía en suministrar al bebé un chupete o tetina que llevaba incorporado un sensor para registrar los movimientos de succión realizados por el bebé.

- Los impulsos recogidos por el sensor eran, a su vez, transmitidos a un amplificador conectado a un ordenador que iba registrando la tasa de succión del bebé ante diferentes estímulos.

- Este procedimiento se basa en la tendencia del bebé a aumentar la tasa de succión con la presentación de un estímulo nuevo, es decir, distinto de otro presentado inmediatamente antes.

- Así, cada vez que un cambio en el TEV producía un aumento en la tasa de succión, esta respuesta se interpretaba **como reflejo de una discriminación perceptiva por parte del bebé;** en caso contrario, es decir si no se producía ningún aumento en la tasa de succión ante la modificación del TEV, se consideraba que el bebé **no había dado una respuesta de discriminación.**

- De esta manera, podía investigarse si las respuestas de discriminación del bebé se ajustaban, al igual que las del adulto, **a cambios categoriales, o si por el contrario se producían también ante modificaciones graduales de la estimulación.**

- En cada ensayo se establecía **una medida de línea base** para un determinado estímulo, el primero del par, es decir, se habituaba al bebé a escuchar repetidamente el mismo estímulo; inicialmente el bebé producía una tasa de succión elevada y creciente, dada la novedad del estímulo y su carácter reforzante; sin embargo, al cabo de varias presentaciones, la tasa de succión decaía por efecto de la habituación.

- **Una vez alcanzada la línea base de respuesta, se procedía a presentar otro estímulo, el segundo del par**, que podía caer dentro o fuera de la categoría fonémica (determinada por el TEV) del estímulo previo. Como puede apreciarse en la figura 8.4, en algunos casos la tasa de respuesta aumentaba considerablemente y se mantenía en valores bastante elevados; en estos casos, el valor de TEV del estímulo "nuevo" caía al otro lado de la frontera crítica "sonoro-sordo" (i.e., +25 mseg.). Sin embargo decaía por debajo de la línea base; en estas ocasiones, el valor de TEV del segundo estímulo del par caía dentro de la misma categoría fonémica que el valor de TEV del primer estímulo.

Por consiguiente**, los resultados de estos experimentos** mostraron que los bebés:

- sólo respondían a modificaciones introducidas en el TEV cuando estas modificaciones suponían un cambio de categoría fonémica,
- es decir, cuando los valores de TEV de cada una de las sílabas del par caían a ambos lados del valor crítico de discriminación (i.e., +25 mseg.), como sucede, por ejemplo, en el par +20/+40.
- Cuando ambos estímulos del para caían a uno u otro lado de esta frontera (vg., 0 y +20, o +40 y +60), no se producían aumentos en la tasa de succión.
- La interpretación de estos resultados parece inequívoca: dado que a tan corta edad los bebés apenas han podido disponer de la experiencia lingüística suficiente para fijar el valor crítico de discriminación de fonemas, esta habilidad debe considerarse innata, es decir, el valor de TEV necesario para discriminar entre consonantes sonoras y sordas se halla genéticamente programado.

En una serie de estudios muy recientes **(Mehler et al., 1988),** se ha podido comprobar que los bebés muestran una sensibilidad muy temprana no sólo hacia estímulos de habla en general, sino especialmente hacia los que corresponden a la lengua que se habla en su comunidad.

- Estudiando a sujetos de tan sólo 4 días de edad, estos investigadores hallaron que bebés pertenecientes a 2 comunidades lingüísticas (francesa y norteamericana) mostraban una tasa significativamente superior de respuestas de succión en presencia de estímulos de habla de su propia lengua frente a la de una lengua "extranjera" (vg., el ruso).

- Según Mehler et al., **los bebés aprovechan fundamentalmente la información prosódica (acento y entonación) para efectuar esta discriminación.**

- Dado que no cabe ninguna posibilidad de experiencia lingüística a tan temprana edad, estos datos constituyen un **indudable apoyo a las hipótesis innatistas del desarrollo del lenguaje**, aunque dejan abierta la posibilidad de un influjo precoz de la experiencia prenatal, un factor difícil de calibrar con los métodos de investigación de que se dispone en la actualidad.

3.4.- Mecanismos de detección de rasgos fonéticos.

La percepción categorial de los sonidos del habla constituye, a primera vista, una prueba de que los procesos de percepción del habla **están basados en la identificación de propiedades fonéticas a partir de claves acústico-articulatorias presentes en la señal**

- Así, por ejemplo, la cualidad de sonoridad de una consonante, es decir, el rasgo que permite discriminar entre fonemas consonánticos sonoros y sordos, **depende del valor de TEV, una propiedad articulatoria que lleva consigo variaciones acústicamente relevantes**.

Según algunos autores, en los procesos de acoplamiento acústico-fonético juegan un papel fundamental lo que se conoce como **"detectores de rasgos",** esto es, sistemas especializados en la detección de propiedades fonéticas a partir de la inspección de fragmentos de la señal de habla.

- Según estas teorías, **los detectores funcionan en paralelo y pueden solaparse unos con otros**. Cada uno tiene un umbral de respuesta que le hace sintonizar de forma variable con los estímulos que recibe, en función de la región en que caiga el estímulo en cuestión.

Veamos un caso hipotético, representado en **la figura 8.5.**

- Supongamos que hay un detector de consonantes sonoras cuya sensibilidad es máxima para los estímulos con un TEV alrededor de 0 mseg., siendo su rango de cobertura el abarcado por la curva de la gráfica.
- En tal caso, la sensibilidad de este detector irá disminuyendo a medida que el TEV de los estímulos se vaya alejando del valor central 0 mseg.

- Este hipotético detector se solapará en los valores superiores de TEV con el detector de consonantes sordas, que tendría su punto de máxima sensibilidad en torno a un TEV de +40 mseg.

- Ello explicaría, por ejemplo, un mayor número de errores de identificación con estímulos próximos a la frontera entre las categorías "sonoro-sordo" (vg., en la región más próxima a +20 mseg.).

Fig 8.5.- Funciones hipotéticas de sensibilidad de los detectores de consonantes sonoras (curvas de la izquierda) y consonantes sordas (curvas de la derecha), sensibles a los valores de TEV correspondientes a los fonemas /b/ y /p/. La gráfica superior refleja la distribución de respuestas perceptivas en condiciones normales. Como puede apreciarse en esta gráfica, el nivel de sensibilidad es máximo en un valor hipotéticamente "central" de TEV para cada categoría (0 mseg., para fonemas sonoros y +40 mseg. Para fonemas sordos) y va disminuyendo progresivamente a medida que nos alejamos de dicho valor y nos aproximamos a la frontera entre las dos categorías. Sin embargo, tras la adaptación por exposición repetida a un fonema sonoro (/ba/) (gráfica inferior), la sensibilidad del detector de fonemas sonoros se reduce, lo que produce un desplazamiento en la frontera entre categorías fonémicas (tomado de Garman, 1990).

1) EXPERIMENTO DE MILLER Y NICELY (1955)

En un experimento ya clásico, **Miller y Nicely (1955)** observaron que **el patrón de confusión en la identificación de sílabas CV se ajustaba al grado de semejanza fonética entre las sílabas "error" y "target"**, esto es, entre la sílaba incorrectamente identificada por el sujeto y la silaba emitida realmente. Así, la probabilidad de cometer un error de identificación era directamente proporcional al número de rasgos distintivos fonéticos compartidos por ambas sílabas.

George A. Miller

- Esto quedaba reflejado en el hecho de que los sujetos tendían a confundir consonantes que eran más semejantes en sus rasgos articulatorios, y que sólo diferían en uno; por ejemplo, /m/ - /n/ , /f/ - /Θ/ , /p/ - /t/ - /k/ , /d/ - /g/, etc., que sólo difieren en el punto de articulación, pero no en el modo de articulación ni en la sonoridad.
- A partir de estos resultados, Miller y Nicely concluyeron que el sistema de percepción del habla **está organizado en canales especializados en la detección de rasgos fonéticos.**

2) EXPERIMENTO DE EIMAS Y CORBIT, (1973)

Una prueba más reciente de la existencia de estos detectores procede del fenómeno conocido como **"adaptación selectiva"** (Eimas y Corbit, 1973).

Este fenómeno se describe como el desplazamiento, en uno u otro sentido, de la frontera entre categorías fonémicas (vg., "sonoro-sordo") por efecto de la exposición repetida al sujeto de un mismo estímulo.

- Así, si a un sujeto se le presenta de forma repetida la sílaba /ba/, al cabo de varias presentaciones se producirá una fatiga del detector de fonemas sonoros y, a consecuencia de ello, una sensibilidad reducida hacia los estímulos de esa categoría.

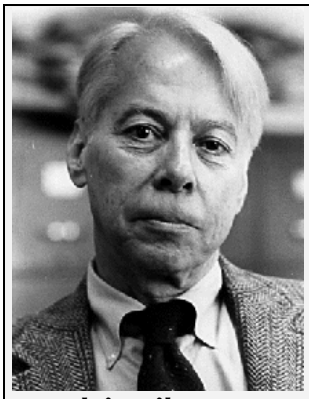
- Por ello, la probabilidad de identificar como sordo (/pa/) un fonema sonoro (/ba/) aumentará a medida que nos acerquemos a la frontera entre las dos categorías fonémicas.
- Ello se debe, según Eimas y Corbit, a que, **por efecto de la fatiga del detector**, se habrá producido un desplazamiento de la frontera entre ambas categorías hacia un valor menor de TEV, según se aprecia en la **figura 8.5** más arriba, en la intersección entre la curva de la izquierda, que corresponde al detector fatigado de fonemas sonoros, y la curva no alterada del detector de fonemas sordos (a la derecha).

Hasta aquí hemos presentado algunas pruebas de **la existencia de mecanismos especializados en la detección de rasgos fonéticos en los procesos de acoplamiento acústico-fonético.**

Sin embargo, a tener de los datos examinados, puede dar la impresión de que existe una correspondencia bastante precisa entre claves acústicas, por un lado, y detectores de rasgos, por otro, de suerte que cada detector cuenta con un número a la vez necesario y suficiente de claves acústicas para realizar su trabajo. Sin embargo, como ya se ha apuntado varias veces, esto no es así.

Algunos investigadores (Liberman, 1982; Eimas, 1985) **han subrayado que la relación entre claves acústicas y representaciones fonéticas es bastante indirecta, y que el proceso de acoplamiento acústico-fonético no es tanto un proceso de "acoplamiento" en sentido estricto, cuanto de "integración".**

Según esto, puede haber claves acústicamente dispares que, en cambio, contribuyan de forma equivalente a producir un mismo efecto perceptivo:


Alvin Liberman

RELACIONES DE INTERCAMBIO ENTRE CLAVES ACÚSTICAS

Liberman (1982) ilustró este fenómeno, conocido con el nombre de **"relaciones de intercambio" entre claves acústicas**, al mostrar la equivalencia perceptiva de dos hechos acústicos tan dispares como pueden ser, por un lado, la transición de los formantes en una sílaba en la que se percibe la consonante oclusiva /p/ (#split#) y por otro, la duración de un intervalo de silencio que antecede a la secuencia /lit/ de dicha sílaba.

- En concreto, la distinción entre la presencia de la oclusiva /p/ (i.e., #split#) y su ausencia (i.e., #slit#) depende, alternativamente, de la manipulación apropiada de las transiciones de los formantes o de la duración de una breve pausa que media entre la consonante fricativa /s/ y la secuencia /lit/.

- Dicho con otras palabras, el espectrograma de #slit# se puede manipular de dos maneras diferentes para inducir el percepto #split#; por un aparte, fijando los valores apropiados de las transiciones de los formantes del fonema /p/, o por otra, aumentando en 20 mseg., la duración del intervalo de silencio que sigue al ruido fricativo correspondiente a la consonante /s/ y antecede a la secuencia /lit/.

- El resultado es, en ambos casos, la sílaba #split#, lo que viene a demostrar **que cada clave acústica no está directamente asociada a una representación fonética individual, sino que todas las claves disponibles tienen que ser integradas de forma apropiada para que surtan un determinado efecto perceptivo**.

- Cómo se explica la integración de claves acústicas en representaciones fonéticas es algo que se expondrá al referirse a las teorías sobre la integración acústico-fonética.

3.5.- Especificidad de la percepción categorial:

Hasta aquí se han examinado algunas pruebas que parecen favorecer el carácter innato de los mecanismos responsables de la percepción categorial del habla. Sin embargo, aún tenemos que responder a los dos interrogantes formulados unas páginas atrás acerca de la especificidad de esta capacidad perceptiva.

Dicha especificidad se refiere,

- por un lado, a la posible aplicación exclusiva de los mecanismos de percepción categorial a estímulos de habla **(especificidad de dominio)** y,
- por otro, al supuesto carácter exclusivamente humano de esta forma de percepción **(especificidad de especie).**

Por lo que se refiere a **la especificidad de dominio** de la percepción categorial, hay que resaltar dos hechos significativos.

a) por una parte, **no todas las distinciones fonémicas relevantes se perciben de manera categorial.**

- Así, por ejemplo, la diferencia entre consonantes fricativas según su punto de articulación (vg., /va/ vs /l↓a/ ; /fa/ vs /Θa/) estriba en la frecuencia relativa de un ruido aperiódico (ruido fricativo) que antecede a la vocal.
- Según parece, las diferencias en el punto de articulación de estas consonantes se perciben, es decir, hay una gradación desde el punto de articulación "labiodental" (/v/, /f/) hasta el "dental" (/l↓/, /Θ/).
- Así pues, no puede afirmarse taxativamente que todas las distinciones acústicas que son fonéticamente relevantes se discriminen de manera categorial, aun cuando buena parte de ellas (vg., las relativas a la sonoridad, ya examinadas, y al punto de articulación en consonantes no fricativas) sí se discriminen de este modo.

b) Un segundo hecho a resaltar es que el fenómeno de **la percepción categorial no sólo se produce en continuos de habla,** como se ha venido señalando hasta aquí, **sino también en continuos de estímulos de no habla**.

- Por ejemplo, en distintos experimentos se ha comprobado que al manipular el desfase temporal existente entre el inicio de dos tonos puros sintetizados, o entre dos estímulos aperiódicos, como un ruido y un zumbido, los sujetos manifiestan respuestas perceptivas categoriales, es decir, **discriminan los dos estímulos del par de forma abrupta y no gradual** (cfr., Miller et al., 1976; Pisoni, 1977; López Bascuas, 1991).
- Y curiosamente, según los datos disponibles hasta ahora, **lo hacen tomando como valor temporal crítico de discriminación el mismo valor de TEV empleado en la discriminación con estímulos de habla en inglés.**

¿Cómo interpretar estos fenómenos?

- Una posibilidad, que se nos antoja bastante improbable, **es que los parámetros de categorización utilizados en la discriminación de estímulos de habla, que parecen ser en buena medida innatos, se impongan asimismo para discriminar estímulos de no habla**, en cuyo caso la percepción auditiva sería tributaria de la percepción del habla.

- Otra posibilidad, sin duda mucho más plausible, es la inversa de la anterior, es decir, **que la percepción categorial es una característica general de la percepción auditiva,** de suerte que los parámetros de discriminación perceptiva del habla (vg., el TEV) toman sus valores de parámetros de la capacidad auditiva general empleados en la resolución de relaciones temporales entre las claves acústicas de diversas clases de estímulos auditivos. Esto supondría que la percepción del habla se halla sometida a procesos aplicados "horizontalmente" a través de distintos dominios estimulares.

Patricia K Kuhl

Algunos datos relativos a **la especificidad de especie de la percepción categorial parecen avalar esta última interpretación.**

En este sentido, **Kuhl y Miller (1978)** comprobaron que la percepción categorial de fonemas oclusivos a partir del TEV no es exclusiva de la especie humana, sino que se manifiesta en otras especies animales, como por ejemplo las chinchillas.

- Utilizando un paradigma de evitación condicionada (a base de emparejar un estímulo de habla con un determinado TEV, vg., +80 mseg., con una descarga eléctrica), **observaron que la respuesta de evitación de la chinchilla a la estimulación era generalizada a todos los estímulos cuyo TEV se situaba aproximadamente por encima de los +23 mseg., pero no a los que caían por debajo de ese valor.**

- En otras palabras, **la chinchilla exhibía una respuesta categorial muy semejante a la expresada por bebés y adultos humanos** (adultos angloparlantes, para ser más exactos) ante la distinción fonémica "sonoro-sordo".

- La conclusión más obvia que cabe extraer de estos resultados es **que la percepción categorial es una capacidad perceptiva no exclusiva de la especie humana** y, por eso mismo, basada en recursos no lingüísticos de procesamiento de la información.

Ahora bien, si, como parece deducirse de las últimas pruebas examinadas, nuestros recursos de procesamiento del habla se apoyan en capacidades auditivas de carácter general,

- **¿cómo se realiza la integración de claves acústicas en la percepción de segmentos fonéticos discretos?,**
- **¿en qué medida puede decirse que, más allá de estas capacidades auditivas generales, existe un mecanismo específico de identificación de sonidos del habla?**

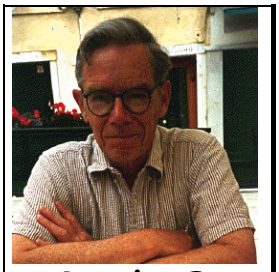
A nadie se le escapa que la mera capacidad para discriminar claves acústicas no equivale a la capacidad para identificar fonemas de la lengua.

Evidentemente, para esto último hace falta algo más. En el próximo apartado se van a exponer dos teorías alternativas acerca de los mecanismos de acoplamiento o

integración de claves acústicas en categorías fonémicas: las teorías “motora” y “auditiva” de la percepción del habla.

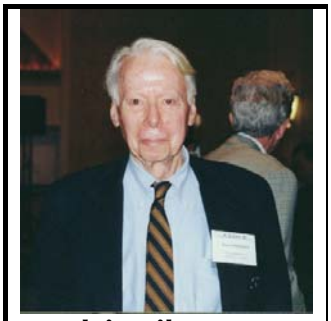
4.- TEORIAS ACERCA DE LA INTEGRACIÓN ACÚSTICO-FONÉTICA.:

1.- Teoría motora de la percepción del habla:



Ignatius G. Mattingly

Una teoría ya clásica de la percepción del habla es la conocida **como “teoría motora”, cuyos más señalado defensores son Alvin Liberman y su grupo de colaboradores en los Laboratorios Haskins (Lieberman et al., 1967; Liberman, 1982; Liberman y Mattingly, 1985).**



Alvin Liberman

Según esta teoría, la identificación fonética se efectúa mediante un sistema de procesamiento especializados en la percepción de sonidos del habla distinto del sistema empleado en la percepción de los restantes estímulos auditivos, lo que determina un modo específico de procesamiento, el llamado “modo del habla” (speech mode).

- El modo de habla se podría definir como un canal de procesamiento de señales auditivas permanentemente sintonizado a aquellas propiedades acústicas de la señal que pueden ser integradas en un estímulo de habla.

- Este sistema se halla adaptado a un código en virtud el cual la estructura fonética del lenguaje se impone sobre las propiedades acústicas de la señal de habla.

- Este código se define, a su vez, en términos de las propiedades articulatorias y coarticulatorias de los sonidos. Es decir, existe un vínculo directo entre los sistemas de percepción y producción del habla que permite al oyente determinar qué gestos articulatorios realiza el hablante y con ello cuáles son los segmentos fonéticos que produce.

- Así pues, la idea central de esta teoría es que el habla se percibe gracias a nuestro conocimiento tácito o inconsciente del modo en que se produce.

- La teoría motora postula como mecanismo básico de la percepción del habla un mecanismo de “análisis por síntesis” que comprende, por un lado, procesos de extracción de información de la señal (análisis), y por otro, procesos de generación “interna” de sonidos a partir de claves acústicas analizadas y del conocimiento de las propiedades articulatorias de los sonidos del habla (síntesis).

- Esto permite explicar cómo se resuelven problemas tan complejos como el de la variabilidad de la señal física del habla (ya aludido) y el del procesamiento de propiedades globales del hablar pertenecientes a niveles superiores de representación.

- La variabilidad se resuelve por medio de la integración de claves acústicas con representaciones articulatorias, una integración que, como ya se ha visto antes, no se produce estrictamente de abajo-arriba, sino de forma interactiva, esto es, acoplando conjuntos de claves acústicas a representaciones fonémicas discretas.

- El mecanismo de análisis por síntesis permite también incorporar a los procesos de percepción del habla propiedades más globales o de orden superior de las emisiones lingüísticas, tales como la estructura suprasegmental (acento y entonación) y métrica (silabificación), dado que estas propiedades pueden influir sobre los procesos de integración acústico-fonética.

PRUEBAS EMPÍRICAS MÁS DESTACABLES A FAVOR DE LA TEORÍA MOTORA:

Son de 3 tipos:

a) En primer lugar, hay evidencia de que **los juicios perceptivos de los oyentes varían sustancialmente en función de información previa sobre las características de la voz del hablante.**

Experimento de Ladefoged y Broadbent (1957), con una tarea de identificación fonética en la que se presentaba auditivamente una breve oración introductoria.

Peter Ladefoged

a) Vg., "Por favor, repita la siguiente palabra", seguida de una secuencia fonética CVC ambigua en cuanto en cuanto a la identidad de la vocal (vg. #b-t#).
b) Los sujetos tenían que repetir en voz alta la palabra fonéticamente ambigua (#b□t#, #b□t#, #b□t# o #b←t#).
c) Las respuestas perceptivas de los sujetos al estímulo crítico (i.e., la palabra que tenían que repetir) variaban en función de las propiedades acústicas de la emisión que antecedía a dicho estímulo cuando ambos (la emisión introductoria y el estímulo crítico) eran producidos por la misma voz.

d) Así, cuando la frase introductoria fue emitida con un timbre elevado, los sujetos identificaron la palabra #b□t# un 87 por ciento de las veces

e) Cuando fue emitida con un timbre bajo, los sujetos identificaron #b□t# en un 90 por ciento de las ocasiones.

f) Esto indica que el sistema perceptivo ajusta sus parámetros a las características acústicas de la fuente de emisión, modelando, por así decir, los gestos articulatorios de dicha fuente (el tracto vocal del hablante).

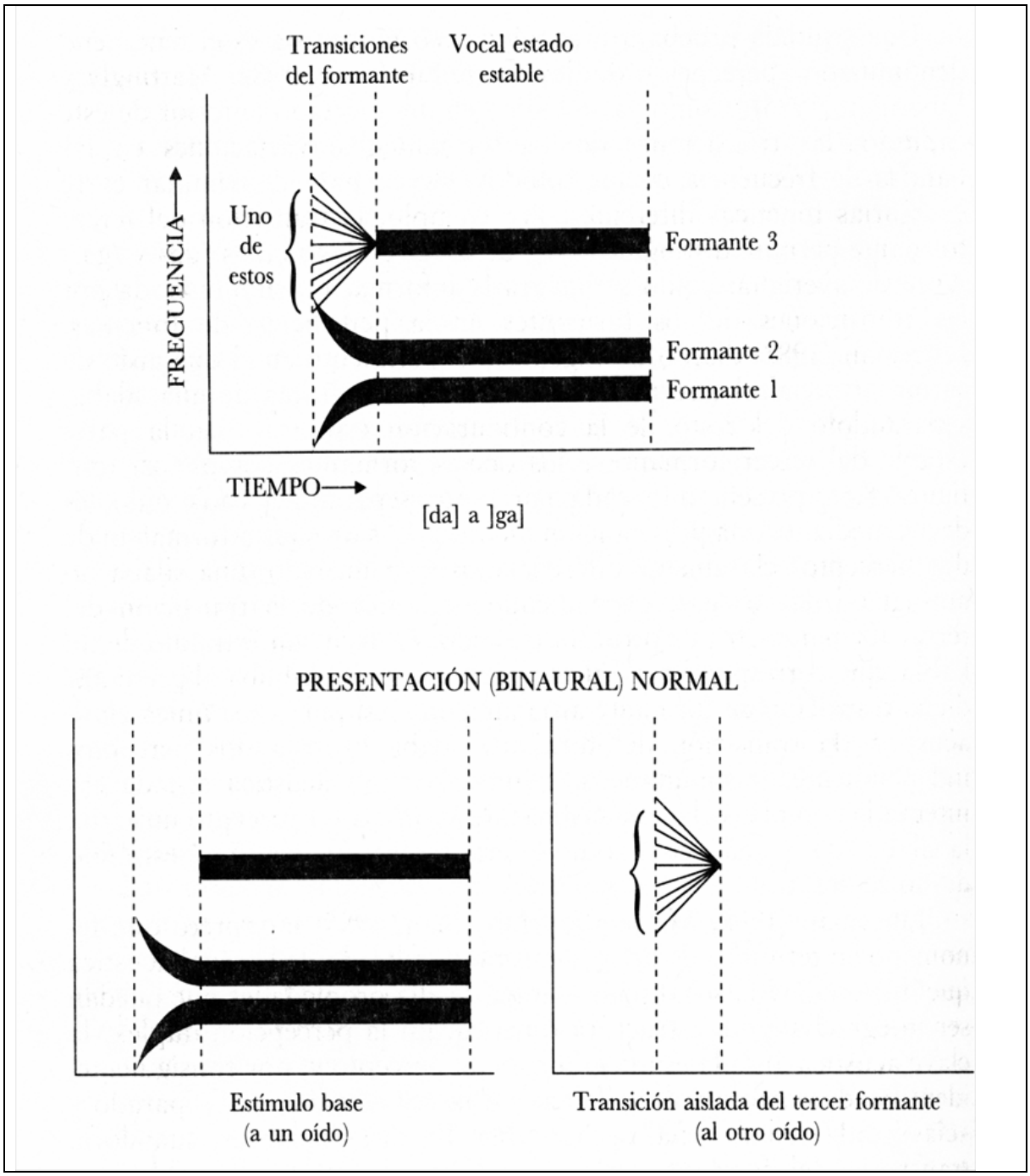
Donald E. Broadbent

b) Una segunda prueba a favor de la teoría motora es **el fenómeno denominado "percepción dúplex"** (cfr., Liberman 1982; Mattingly y Liberman, 1985). Como ya se ha indicado, las transiciones de los formantes, o variaciones en las bandas de frecuencia de los sonidos, **sirven para discriminar entre categorías fonéticas diferentes.** Por ejemplo, la transición del tercer formante permite distinguir la /d/ de la /g/ en las sílabas /da/ y /ga/.

A fin de averiguar cómo se integra la información suministrada por las transiciones de los formantes en la percepción de fonemas, Liberman (1982) diseñó un ingeniosos experimento en el que aisló de forma artificial el tercer formante del espectograma de una sílaba separándolo del resto de la configuración espectral (i.e., la parte estable del tercer formante y los demás formantes completos) (cfr. Figura 8.6), presentando cada parte por separado a cada oído, es decir, mediante una presentación dicótica.

a) Los oyentes informaban de dos perceptos claramente diferenciados: por una lado,
   - **una sílaba no ambigua** (/da/ o /ga/, dependiendo del valor de la transición del tercer formante),
   - y por otro, **un chirrido**, es decir, un estímulo de no habla, que correspondía al mismo percepto que resultaba al presentar dicha transición de formante aisladamente.
b) Así pues, una única clave acústica **(la transición del formante)** daba lugar a dos perceptos independientes y simultáneos: cuando la calve acústica aislada era integra con otras claves acústicas, se forma un percepto unitario: la sílaba /da/ o /ga/; cuando no lo era, se percibía como un estímulo de no habla.
c) Liberman (1982; Mattingly y Lieberman, 1985) interpretó este fenómeno en términos de **un mecanismo de filtrado de la señal acústica que inspecciona la señal para extraer de ella propiedades que puedan ser integradas en la estructura fonética.**
d) **En la percepción dúplex**, la clave acústica aislada es integrada en un percepto y, por consiguiente, identificada como señal de habla, y a la vez recogida por separado y "clasificada" como señal de no habla.

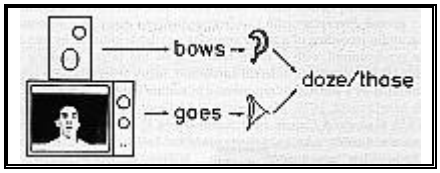
e) En otras palabras, cuando la transición del formante es percibida en el "modo de habla", es aprovechada para identificar fonemas; cuando no lo es, es objeto de un análisis acústico en el sistema auditivo general.



c) La tercera prueba que se presenta a favor de la teoría motora **procede de estudios sobre integración de información visual y auditiva en la percepción de sonidos de habla.** Si la teoría motora es correcta, cualquier información sobre las propiedades articulatorias del habla que pueda ser empleada en la identificación de sonidos será utilizada por los oyentes y, por tanto, influirá en los juicios y respuestas perceptivas de los sujetos.

a) un dato revelador en este sentido es la observación de que cuando a bebés de pocos meses de edad se les presentan estímulos de habla (vg., vocales) emparejados con imágenes de un hablante efectuando gestos articulatorios que bien corresponden o bien contradicen el estímulo auditivo, **los bebés muestran un mayor interés (medido a través del tiempo de contemplación del estímulo visual)** por la estimulación visual cuando ésta coincide con el estímulo percibido auditivamente (cfr. Miller, 1990).
   - Es decir, los bebés pasan más tiempo contemplando un rostro que está articulando la vocal /i/ cuando simultáneamente escuchan el fonema /I/.
   - Esto sugiere que **los bebés poseen un conocimiento tácito de la relación entre las consecuencias auditivas y visuales de la articulación de fonemas.**

b) Otra prueba en este sentido es el llamado **"efecto de McGurk"** (McGurk y MacDonald, 1976).
   - Estos investigadores hallaron que cuando a un sujeto perceptor se le presentan estímulos auditivos y visuales contradictorios en cuanto a sus consecuencias fonéticas, los sujetos adoptan de forma inconsciente una solución de compromiso entre ambas fuentes de estimulación.



   - Por ejemplo, si el estímulo auditivo es una consonante con un punto de articulación anterior (vg., la consonante bilabial /ba/) y el estímulo visual muestra a un hablante articulando una consonante posterior (vg., la consonante velar /ga/), el percepto resultante es una consonante dental /da/, cuyo punto de articulación se encuentra a medio camino entre los de las consonantes percibidas por cada modalidad sensorial.
   - Así pues, lejos de producirse un conflicto "bimodal" (entre las modalidades sensoriales visual y auditiva), se origina un percepto que no es ni puramente visual ni puramente auditivo, sino fonético-articulatorio.

2. Teoría auditiva de la percepción del habla:

Según esta teoría, la percepción del habla no requiere ningún sistema especializado de procesamiento, sino que el habla se percibe por medio de los mismos mecanismos que cualquier otro estímulo auditivo.

- A diferencia de la teoría motora, **la teoría auditiva no es un paradigma unificado de explicación, sino que reúne una variedad de modelos y explicaciones en muchos aspectos divergentes entre sí.**
- En términos generales, sin embargo, los enfoques "auditivos" de la percepción del habla **tienden a rechazar la idea de que la percepción del habla sea específica en los dos sentidos arriba apuntados, es decir, específica de dominio y específica de especie.**
- Dentro de este marco teórico general**, se rechazan las explicaciones basadas en mecanismos de análisis por síntesis a favor de mecanismos en general más analíticos.**
- Así, los partidarios de este tipo de explicaciones **tienden a investigar los mecanismos de procesamiento temprano de la señal, situados en los niveles auditivos de análisis.**

Esto puede traer consigo 2 consecuencias diferentes:

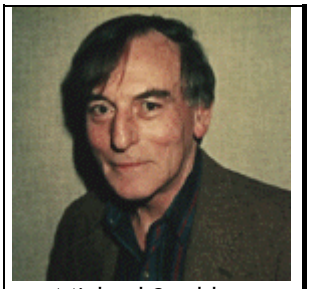
a) por una parte, hay autores para **quienes la señal de habla no es tan variable como se supone**, sino que en el nivel acústico existen propiedades invariantes que permiten efectuar un acoplamiento microestructural directo, por llamarlo así, entre las propiedades físicas y las representaciones fonéticas (cfr., Blumstein y Stevens, 1979, 1989).

- Una clave acústica que parece estar directamente asociada con propiedades articulatorias de los fonemas, en concreto con el punto de articulación de fonemas oclusivos (vg., bilabiales como /b/ o /p/, alveolares como /d/ o /t/, o velares como /g/ o /k/), es la transición del segundo formante. Este acoplamiento microestructural, al que ya se ha hecho referencia, tiene lugar en la etapa de análisis auditivo central (cfr., punto 3.1).

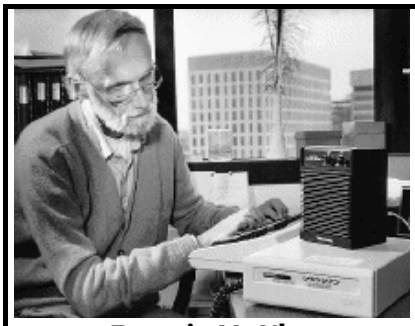
Sheila E. Blumstein

b) En cambio, otros autores (vg., Klatt, 1979, 1980; Studder-Kennedy, 1981, citado en Garman, 1990) sostienen que **las propiedades invariantes de la señal acústica no emergen en el nivel microestructural, sino al contrario, en niveles macroestructurales, y en concreto, en el nivel léxico**.

- Desde este punto de vista, que para distinguirlo del anterior se denominará acoplamiento macroestructural, se mantiene que la señal de habla estimula unos "patrones neurosensitivos" que representan formas léxicas almacenadas en la memoria (Studdert-Kennedy, 1981)


Michael Studdert-Kennedy

c) En la versión de Klatt (1979), **estas representaciones léxicas están formadas por "plantillas espectrales", o secuencias de representaciones espectrales prototípicas.**



Dennis H. Klatt

- Una plantilla espectral es, pues, **la representación de una secuencia "ideal" de claves acústicas.**
- A medida que el oyente recibe fragmentos de la cadena hablada, va computando sobre la marcha representaciones espectrales de la señal (que Klatt denomina **"difonos"**) y comparando estas representaciones con las que tiene almacenadas en la memoria (las plantillas espectrales).

- La característica fundamental tanto de las plantillas como de los difonos es que **contienen información sensible al contexto acústico en el que se producen, esto es, son representaciones dependientes del contexto.**
- Esto **permite resolver el problema de la ausencia de invarianza en el plano acústico.**
- Sin embargo, estas representaciones **no corresponden a unidades fonémicas discretas;** es decir, no existe en este modelo ningún nivel de representación fonética o fonológica, sino que del nivel acústico se accede directamente al nivel léxico.
- Por ello, al modelo de Klatt de reconocimiento de palabras, basado en un acoplamiento macroestructural, se le conoce como **"acceso al léxico a partir de espectro" (Lexical Access From Spectra, o LAFS)**

Se señalaba al comienzo del capítulo como dos problemas fundamentales de la percepción del habla **el de la segmentación fonémica** y **el de la ausencia de invarianza.**

Sin embargo, si se considera el problema de la percepción del habla en un contexto más amplio, es decir, si se tiene en cuenta que forma parte de un conjunto de procesos orientados a la comprensión de mensajes lingüísticos con significado, habrá que prestar atención a otros factores de orden superior que también pueden influir en las primeras etapas de la identificación de sonidos del habla.

Nos estamos refiriendo a factores lingüísticos y contextuales tales como:

- **las restricciones suprasegmentales,**
- **el conocimiento léxico, e incluso**
- **variables sintácticas y semánticas**.

Aunque la discusión de esta clase de factores tal vez resulte más pertinente en otros niveles de procesamiento, se va a hacer ahora una breve referencia al papel que pueden desempeñar en los últimos niveles de la percepción del habla, esto es, en las etapas fonética y fonológica.

5.- LA PERCEPCION DEL HABLA CONTINUA:

El título de este apartado puede parecer equívoco por lo que tiene de tautológico. A causa de las restricciones impuestas por la coarticulación, **el habla no se puede producir segmento a segmento, sino únicamente de forma continua.**

- Precisamente en eso reside una de las paradojas de la percepción del habla, en el hecho de que siendo el habla humana un estímulo que se despliega de forma continua, tenga que pasar por un estadio en el que se identifican segmentos discretos.

- Sin embargo, a lo que habitualmente se alude cuando se hace referencia a la "percepción del habla continua" es al hecho de que **la percepción del habla humana es un proceso activo cuyo funcionamiento viene determinado no sólo por información de la señal física, sino también por información lingüística y extralingüística que el oyente tiene almacenada en su memoria.**

- Dicho de otro modo, los sonidos del habla **se encuentran inmersos en un contexto lingüístico (sílabas, palabras, oraciones) que el sujeto perceptor analiza e interpreta mediante procesos que operan de forma concurrente con los procesos de identificación de sonidos.**
- Por consiguiente, es posible que los procesos psicolingüísticos más inmediatos a los de la percepción del habla (vg., reconocimiento de palabras, análisis sintáctico o comprensión del significado oracional) ejerzan un influjo descendente o de "arriba –abajo" sobre ellos.
- Una consecuencia bastante obvia del carácter activo de la percepción del habla continua es que **el procesamiento de la señal acústica no tiene por qué ser exhaustivo, es decir, no es preciso identificar todos y cada uno de los segmentos fonémicos de la entra sensorial para acceder a otros niveles superiores de procesamiento, tales como el reconocimiento de palabras.**
- En consecuencia, puede decirse que, al menos a partir de los niveles de análisis fonético y fonológico, **se produce una interacción entre procesos de identificación de segmentos fonéticos a partir de claves acústicas, por una lado, y procesos de acceso al léxico a partir de representaciones fonológicas, por otro.**

Esta conexión entre procesos fonológicos y léxicos será objeto de reflexión en el próximo capítulo, cuando se examine la naturaleza de las representaciones que el oyente o el lector utiliza para acceder a su léxico mental y reconocer palabras.

Veamos algunos ejemplos de la interacción entre procesos perceptivos y procesos lingüísticos de orden superior.

- El conocimiento de ciertas peculiaridades prosódicas y fonológicas de las secuencias fonéticas del castellano, como por ejemplo **el que todas las palabras de contenido tengan una sílaba tónica o acentuada, o el que los grupos consonánticos iniciales de sílaba sólo puedan llevar como consonante "trabada" una líquida**, /l/ o /r/ (vg., #pla#, #bra#, #cla# o #gla#, pero no #pna#, #bsa# o #cta#), **puede imponer restricciones sobre los procesos de acoplamiento acústico-fonético**, permitiendo al oyente anticipar eventos acústicos a partir de propiedades prosódicas y fonológicas del contexto.

Una de las motivaciones de la teoría motora de la percepción del habla es probablemente **la necesidad de incorporar este tipo de conocimientos a las primeras etapas de los procesos perceptivos**.

Sin embargo, **es muy posible que información de orden todavía superior (vg., léxica, sintáctica o semántica) intervenga también en los procesos más primarios de análisis de la señal.**

- Por ejemplo, en un experimento ya clásico (Pollack y Pickett, 1964, comentado en Clark y Clark, 1977) se presentaba a los oyentes fragmentos de una conversación, consistentes en palabras completas extraídas de su contexto oracional, y se les pedía que identificaran esas palabras.

- El porcentaje de respuestas correctas apenas sobrepasó el 50 por ciento, lo que se interpretó como indicio de que la señal de habla permanece ininteligible hasta que el perceptor no dispone de suficiente información contextual para formular hipótesis sobre el contenido fonético del mensaje.

- **Esta información contextual incluye información no sólo léxica, sino también supraléxica (prosódica, sintáctica y semántica), dado que de no ser así los sujetos hubieran podido identificar fonemas a partir de las palabras sueltas suministradas en la grabación, cosa que no fueron capaces de hacer.**

A pesar de sus llamativos resultados, el experimento de Pollack y Pickett no puede, sin embargo, aducirse como prueba contundente a favor del carácter activo de los procesos de percepción del habla. **Esto se debe a que el juicio que se les pedía a los sujetos era un juicio de reconocimiento de palabras, y no un juicio estrictamente perceptivo.**

Por consiguiente, se puede argüir **que el nivel de procesamiento involucrado en esta tarea incluye otros procesos además de los propiamente destinados a la identificación de fonemas.**

La necesidad de evitar cualquier contaminación de los procesos perceptivos más periféricos por procesos de niveles superiores es lo que explica que los experimentos de identificación y discriminación perceptiva se efectúen con estímulos acústicos desprovistos de significado lingüístico (vg., sílabas o secuencias de fonemas en lugar de palabras).

Las pruebas sin duda más fehacientes del influjo de procesos superiores de reconocimiento sobre mecanismos perceptivos elementales proceden **del estudio de fenómenos que caen dentro de la categoría de las "ilusiones perceptivas**". Nos vamos a referir a 3 de estos fenómenos:

1) **La "restauración de fonemas".**
2) **El "seguimiento con restauración de errores".**
3) **La "escucha selectiva".**

1.- El efecto de restauración de fonemas (cfr., Warren, 1970)

Consiste en la sustitución inconsciente de material fonético, ausente de la señal acústica, por un estímulo de no habla presente en la señal. A los sujeto se les presentaban auditivamente oraciones como:

*1.- Los gobernadores de los estados se reunieron con sus respectivas leg*laturas antes de la convocatoria del pleno del congreso.*

- A las que se les había extraído una porción de la señal (de 120 mseg., de duración) que correspondía aproximadamente a un fonema (/s/ en el ejemplo), sustituyéndola por un sonido de no habla (vg., una tos o un zumbido); más tarde, se les preguntaba si notaban la falta de algún sonido y dónde localizaban el estímulo de no habla.

- La mayoría de los sujetos afirmaron haber percibido la palabra crítica "legislatura" intacta, o sea, con todos sus fonemas, y la mitad de ellos localizaron el sonido de no habla en una de las fronteras de la palabra.

- Este efecto persistía incluso al repetir la grabación y aunque el sujeto supiera que en realidad el fonema /s/ había sido extraído de ella.

- El efecto, sin embargo, desaparecía cuando el fonema ausente pertenecía a una pseudopalabra (vg., "apismatura").

- En posteriores estudios sobre este mismo efecto, se pudo comprobar que la restauración también se efectúa cuando el fragmento sustituido abarca más de un fonema (vg., "le***latura") e incluso se halla condicionada a los contextos sintáctico y semántico en que se encuentra la palabra crítica.

Así, cuando la ausencia de información acústica origina una secuencia ambigua desde el punto de vista léxico (vg., "ca*a", donde la restauración de un fonema u otro puede dar lugar a diferentes palabras: "casa", "cama", "cara", etc.), y además se dispone de un contexto oracional desambiguante (vg., "la ca*a no tenía puertas"), el efecto de restauración se acomoda a una palabra semánticamente congruente con el contexto oracional inmediato ("casa" en el ejemplo anterior).

2.- El efecto de restauración de errores.

Un fenómeno muy similar a la restauración de fonemas es el efecto de restauración de errores, en el que el sujeto, también de forma inconsciente, sustituye estímulos de habla erróneos (vg., errores fonéticos como "cerveza") por las formas correctas ("cerveza").

Este efecto se registra con bastante consistencia **con el uso de la llamada tarea de "seguimiento" (shadowing), en la que el oyente recibe un mensaje verbal por vía auditiva y tiene que repetirlo en voz alta de forma simultánea a su recepción y a la mayor brevedad posible.**

William Marslen-Wilson

- **Marslen-Wilson (1973, 1975)** observó que la tarea de seguimiento se puede realizar con latencias muy breves (de hasta 250 mseg., o una sola sílaba de desfase entre el mensaje de entrada y el mensaje de salida), y que las correcciones efectuadas por los sujetos son sensibles a información no sólo léxica, como puso de manifiesto el efecto de restauración de fonemas, sino también sintáctica y semántica.

- Así, los sujetos no sólo corrigen sobre la marcha errores fonéticos, que afectan a sonidos aislados, sino también errores sintácticos y semánticos, sustituyendo palabras erróneas por otras adecuadas al contexto de la oración en que aparecen (vg., "carta" por "hacia", en una oración como "La secretaria introdujo el papel en la máquina y se dispuso a escribir la hacia").

Estos hallazgos ponen de relieve que la información lingüística de orden superior se halla disponible desde las primeras etapas del reconocimiento y puede ejercer un influjo considerable sobre los procesos de percepción del habla continua.

3.- El fenómeno de "Escucha selectiva"

Una última prueba que revela el empleo de información lingüística de orden superior en los procesos de identificación del habla es el fenómeno de "escucha selectiva" (cfr. Cherry, 1953, Treisman, 1964, citados en Clark y Clark, 1977).

Anne Treisman

- Cuando a un oyente se le somete a una tarea dicótica, instruyéndole a que efectúe un seguimiento de uno de los canales, haciendo caso omiso del otro, se registran una serie de efectos originados por el material presentado por el canal no atendido.

- En principio, se comprobó **que si bien el oyente advierte ciertos cambios en el mensaje no atendido** (vg., el cambio de una voz femenina por otra masculina, de estímulos de habla por estímulos de no habla, o de habla normal por habla hacia atrás), **no es capaz de recordar el significado de dicho mensaje ni siquiera de reconocer palabras presentadas repetidamente en el mismo.**

- Sin embargo, en determinadas circunstancias, **el mensaje no atendido puede provocar interferencias en la tarea de seguimiento cuando comparte ciertas características con el mensaje atendido** (vg., cuando contiene semejanzas acústicas, léxicas o semánticas).

Asimismo, cierta información del mensaje no atendido puede influir sobre el procesamiento del mensaje atendido.

- Por ejemplo, Lackner y Garrett (1972) hallaron que **cuando un mensaje atendido contiene una palabra ambigua y el no atendido presenta una palabra asociada** con uno de los significados de la palabra ambigua, el oyente hace uso de esta información a la hora de interpretar el mensaje presentado por el canal atendido.

James R. Lackner

- Así, en una oración como "El hombre divisó la presa a lo lejos", la palabra "presa" se interpretaba en su acepción de "víctima" cuando en el canal no atendido se presentaba la palabra "animal", y en su acepción de "embalse"cuando la palabra no atendida era "agua".

¿Qué conclusiones cabe extraer de estos fenómenos?

- En una primera aproximación, se puede concluir que los procesos de reconocimiento del habla están abiertos a influencias de niveles superiores de procesamiento, es decir, que la información suprasegmental (prosódica y métrica), léxica, sintáctica y semántica impone restricciones sobre los procesos de percepción del habla continua, y que estas restricciones operan, en gran medida, de forma automática e inconsciente.

- No obstante, de ello no se puede concluir, como pretenden algunos autores, que la percepción del habla continua sea un proceso completamente distinto de la percepción de sonidos aislados.

- Más bien se debe subrayar que la percepción del habla es un proceso sujeto a dos tipos de determinantes:

- por un lado, **restricciones impuestas por las propiedades físicas de la señal** y,
- por otro lado, **restricciones impuestas por las representaciones lingüísticas** que se recuperan en niveles superiores de procesamiento.

Es indudable que ambas clases de información convergen e interactúan en algún punto. Lo difícil es establecer a partir de qué momento del procesamiento de la señal comienzan a ejercer sus efectos los procesos de identificación de unidades lingüísticas de carácter superior (vg., palabras y oraciones).

Los procesos de reconocimiento de palabras podrían ser un terreno adecuado para ello.

6.- PERCEPCION DEL LENGUAJE ESCRITO (OPTATIVO)

Las actividades psicológicas reunidas bajo el concepto de "lectura" (lo que, en términos más técnicos, se conoce como **"procesamiento visual del lenguaje"**), encierran un vastísimo repertorio de habilidades, procesos y componentes de información que apenas se podrían exponer con suficiente profundidad en los límites relativamente estrechos de un texto como el presente.

Pese a que la lectura es uno de los temas clásicos con "carta de ciudadanía" en psicología del lenguaje, hasta el punto de que desborda los límites de esta disciplina para adentrarse en otras fronterizas como la psicología de la memoria o de la percepción, en el campo de las disciplinas básicas, o como la psicología de la instrucción, en el de las aplicadas, no hemos considerado oportuno dedicarle un capítulo independiente en este libro, pues ni siquiera eso bastaría para cubrir adecuadamente tan amplios y variados contenidos.

Se ha optado, en cambio, por analizar los componentes del procesamiento del lenguaje escrito en sus distintos niveles asignando cada uno de ellos a la categoría de los procesos psicolingüísticos que lo engloba y con ello al capítulo correspondiente de este libro, en aras de mantener la unidad temática y lógica de exposición de los procesos manifestada en el capítulo 7, que sirvió de introducción a los procesos de actuación lingüística.

Así pues, se comenzará por examinar en este apartado los componentes básicos de los procesos de lectura, ubicados en el nivel perceptivo, para tratar en el capítulo 9 el reconocimiento visual de palabras, en el 10 la comprensión de oraciones escritas, y en el 11 la comprensión de textos.

Con todo, hay que hacer notar que no siempre se va a mantener una nítida división entre el procesamiento del lenguaje en las modalidades visual y auditiva, ya que a medida que nos movemos hacia niveles más centrales de procesamiento, esa distinción encuentra cada vez menos justificación empírica y las coincidencias entre la comprensión del lenguaje oral y la comprensión del lenguaje escrito superan cada vez más a las diferencias.

Comenzando, pues, por el estrato más básico de la percepción visual del lenguaje,

nos encontramos con dos problemas fundamentales:

1) por una parte, es importante determinar
   - Cuáles son **los niveles de representación de la información visual** que median entre la señal estimular y el léxico, y en concreto, la relación que hay entre los procesos de identificación de los rasgos visuales de las letras (considerados las unidades mínimas de codificación de la información visual, que desempeñan un papel similar al de los rasgos distintivos de los fonemas en el lenguaje oral),
   - **Los procesos de identificación de letras** y
   - **Los procesos de identificación de palabras**;
2) y por otra parte, es preciso conocer **los mecanismos a través de los cuales se extrae la información visual del texto durante la lectura y el funcionamiento de estos mecanismos;** a este respecto, nos centraremos en el estudio de las pautas de movimientos y fijaciones oculares que realizan los lectores experimentados.

6.1.- Niveles de procesamiento en la lectura y sus relaciones:

El lenguaje escrito, lo mismo que el oral, consta de diversas clases de unidades que se superponen jerárquicamente, de suerte que las unidades de mayor tamaño resultan de la combinación recursiva de unidades más pequeñas.

- Así, los textos, unidades mayores, se componen de **oraciones**; éstas, a su vez, están formadas por una sucesión de una o más **cláusulas**; las cláusulas están integradas por **palabras** y las palabras son secuencias de **letras**. Inclusive las letras están constituidas por unidades más pequeñas, como son **los trazos o rasgos visuales** (vg., líneas verticales, horizontales u oblicuas, trazos circulares, etc.).

Atendiendo a esta característica, **la lectura puede definirse** *como un proceso de identificación de estas diversas clases de unidades y de integración de las mismas en unidades superiores, y, ocasionalmente a la inversa, como un proceso de descomposición de unidades mayores (vg., palabras) en unidades más elementales (vg., letras).*

Una idea intuitivamente plausible, que además ha recibido abundante confirmación empírica, es que **las palabras son unidades "privilegiadas" de identificación en el transcurso de la lectura.**

- Esto equivale a decir que las unidades léxicas del texto han de ser necesariamente descifradas o identificadas para poder comprender el lenguaje escrito, sin que sea imprescindible, en cambio, la identificación exhaustiva de unidades inferiores a la palabra.

A esta hipótesis se le puede agregar el supuesto de que **las unidades "palabra" se identifican, de ordinario, con mayor rapidez que las letras o sílabas** que forman parte de ellas, pese a que estas últimas son unidades más pequeñas y jerárquicamente inferiores aquéllas, es decir, el acceso a la palabra se produce antes que el acceso a las partes integrantes de la misma.

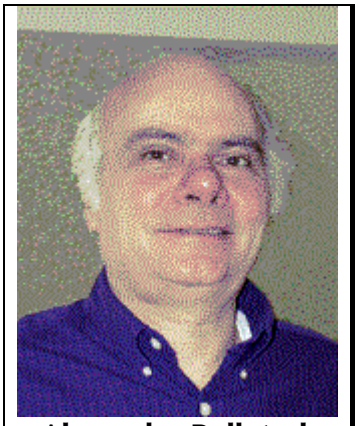
- A esta hipótesis se la conoce también con el nombre de **"modelo de reconocimiento directo o global de palabras",** y mantiene que los procesos de reconocimiento de letras y rasgos son independientes de, amén de innecesarios para, los procesos de identificación de palabras, toda vez que éstas se suelen procesar como una configuración estimular global (cfr. Smith, 1971).

- Esta hipótesis es compatible con la observación que se ilustrará en el próximo capítulo, de que las palabras (tanto orales como escritas) se reconocen a menudo antes de ser percibidas (o presentadas) en su totalidad, o sea, que no es necesario analizar la palabra en todos sus componentes elementales (sílabas, letras) para poder identificarla.

El efecto experimental sin duda más contundente a favor del modelo de reconocimiento global de palabras es el llamado **"efecto de superioridad de la palabra",** descubierto hace ya casi un siglo (Cattell, 1886 o Pillsbury, 1897, citados en Cuetos, 1990) y replicado en numerosas ocasiones (Reicher, 1969; Johnston y McClelland, 1974).

- Según este efecto, **la latencia de reconocimiento de una letra** (vg., la letra "L") **es más breve, y superior el porcentaje de identificaciones correctas de la misma**, cuando esa letra forma parte de una palabra (vg., ABRIL) que cuando forma parte de una no-palabra (vg., IRBAL) o incluso cuando se presenta asiladamente (en el mismo punto del espacio que ocuparía si formara parte de una palabra, vg., _____L) (Reicher, 1969).

¿Cómo puede explicarse este efecto?


Alexander Pollatsek

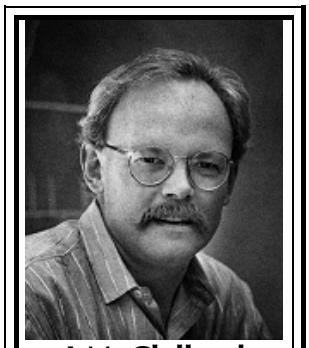
Según Pollatsek y Rayner (1989), caben dos interpretaciones alternativas:

a) una en términos estrictamente "de arriba-abajo" y
b) otra de carácter interactivo.

1) La primera interpretación **afirma que el sistema de reconocimiento dispone de unas "plantillas" o representaciones internas globales de las palabras a las que se acopla la entrada sensorial, sin que haya intervención de representaciones intermedias (vg., de rasgos y/o letras.**

2) **<u>La interpretación interactiva</u>**, por su parte, *explica el efecto de superioridad de la palabra apelando a un procesamiento en paralelo de letras y palabras, de tal modo que la identificación de una palabra facilita la identificación de las letras que la componen.*

- Este interpretación ha sido formulada **en términos de un modelo conexionista o de activación interactiva** (McClelland y Rumelhart, 1981; McClelland, 1987), en el que existe un flujo de activación multidireccional entre los niveles de las palabras, las letras y los rasgos visuales, según se ilustra en **la figura 8.7.**


J McClelland

David Rumelhart

- A raíz de la presentación de un estímulo visual, se produce una activación en cascada de los nodos de rasgos, letras y palabra que configura un patrón de activación global en el sistema.
- Las unidades que alcanzan niveles superiores de activación son "seleccionadas" para la decisión perceptiva, es decir, para la identificación.
- Así, puede suceder que una palabra alcance el nivel requerido de activación antes que algunas de las letras que la integran, o que una determinada letra consiga sumar más activación (por efecto del flujo descendente procedente del nivel de la palabra) cuando se halla incluida en una palabra que cuando está aislada o pertenece a una no-palabra (que, como es lógico, carece de representación en la red de nodos).

Los modelos de plantillas han sido reiteradamente criticados en psicología de la percepción por su incapacidad para librarse de los rasgos más superficiales e irrelevantes de los estímulos y, en consecuencia, para dar cuenta del uso de propiedades abstractas de los mismos en el proceso de categorización perceptiva.

- Parece incuestionable que **nuestra memoria no puede almacenar tantas plantillas** correspondientes a una determinada palabra como formas visuales distintas pueda adoptar (¿algún lector tiene dificultades para descifrar algunas de estas "versiones" de la misma palabra?: ABRIL, abril, aBrIl, AbRiL).

Kathryn Spoehr

- Pero aparte de esta insuficiencia, hay un fenómenos muy similar al efecto de superioridad de la palabra que los modelos de comparación de plantillas se muestran, por razones obvias, incapaces de explicar y que, en cambio, es perfectamente abordable por un modelo interactivo, a saber, **la superioridad de las pseudo-palabras tanto sobre las letras aisladas** (Spoehr y Smith, 1975) como, en ocasiones, sobre las propias palabras (Baron y Thurston, 1973).

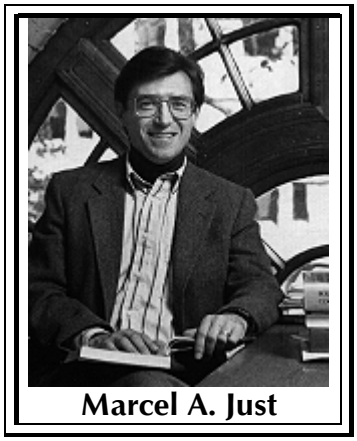
<u>La explicación de este efecto</u> en términos interactivos radica en la idea de que los nodos de letra activados al presentar una pseudo-palabra (vg., IRBAL) propagan activación hacia todas las palabras que contienen dichas letras (vg., el nodo "L" propagará activación de intensidad variable hacia los nodos de palabra "METAL", "CORDEL", "AVAL", entre otros muchos); éstas, a su vez, devuelven la activación hacia los nodos de letra, contribuyendo así a incrementarla (**Pollatsek y Rayner, 1989**; y Cuetos, 1990, para un desarrollo más completo de esta explicación).

Keith Rayner

Por supuesto, para que esta explicación sea plenamente convincente, habrá que tener en cuenta **no sólo la magnitud de la activación** que alcanzan los nodos de la red en un momento dado, sino también **la dimensión temporal** en la que acontecen los procesos de propagación de la activación.

Según se puede colegir de lo expuesto hasta aquí, las letras (y probablemente sus rasgos visuales) son unidades que desempeñan un papel específico y relevante en el procesamiento lector, aunque su identificación sea algo más "soterrada" o se vea desplazada por la "sapiencia" perceptiva de las palabras. Parece, inclusive, que hay razones para creer que el proceso de identificación de rasgos y letras se inicia antes (aun cuando concluya después) que el de identificación de palabras.

En lo que concierne al carácter serial o paralelo de los procesos de identificación de letras en palabras, los datos son, por el momento contradictorios.
La evidencia más concluyente **a favor de la identificación serial de las letras procede de estudios cronométricos de la duración de las fijaciones oculares durante la lectura.**


Marcel A. Just

- Según han comprobado **Just y Carpenter** (1987), las fijaciones oculares presentan una duración proporcional a la longitud, medida en número de letras, de la palabra sobre la que se fija la mirada, cosa que no tendría por qué ocurrir si las letras fueran procesadas en paralelo.

- Frente a estos resultados, **hay pruebas de que los mecanismos de identificación de letras pueden abarcar varias letras de una vez**. Así lo confirma, por ejemplo, la observación de que las palabras con caja alterna entre letras mayúsculas y minúsculas (vg., aBrIl) se leen más despacio que las de caja uniforme (vg., abril o ABRIL), o el hecho, antes aludido, de que las palabras se suelen identificar sin necesidad de analizar todas sus letras componentes.

- Un dato revelador a este respecto es el llamado **"efecto de invisibilidad de las letras",** que se traduce en la comisión de errores de omisión en tareas de detección de letras en contexto.

- Según se ha comprobado (cfr., Carroll, 1986), **la probabilidad de errar en la detección de una letra es directamente proporcional a la frecuencia de uso de la palabra a la que pertenece dicha letra y depende**, asimismo, de factores tales como:

D. W. Carroll

- **La posición que ocupa la letra en la palabra** (siendo más fácil pasar por alto letras situadas en la porción final de las palabras) y de
- **La clase de vocabulario a la que pertenece la palabra** (se detectan peor las letras en palabras funcionales o de clase cerrada (vg., artículos, preposiciones, conjunciones) que en palabras de contenido o clase abierta (vg., nombres, verbos, adjetivos).

Como se ha comprobado en este mismo capítulo al tratar de la percepción del habla, la cuestión de la naturaleza de las unidades de procesamiento (rasgos vs fonemas o letras vs palabras, en la percepción; sílabas vs palabras, en el reconocimiento léxico; o claúsulas vs proposiciones, en la comprensión de oraciones) es un debate eternamente inconcluso.

En gran medida ello se debe a que la elección de las unidades óptimas de procesamiento en tareas de actuación lingüística es un fenómeno sujeto a numerosos factores condicionante. En el caso particular de la lectura, se han señalado 4 (cfr., Vellutino, 1982, citado en Cuetos, 1990):

1) **<u>El contexto</u> en el que se halla la palabra**: las palabras presentadas en un contexto lingüístico se hallan sometidas a restricciones de orden superior (i.e., sintácticas y semánticas), en tanto que las palabras aisladas no. La disponibilidad de esta información "supraléxica" puede facilitar el empleo de un análisis global de la palabra, al suprimir de antemano muchos de los posibles candidatos léxicos.
2) **<u>Ciertas características de la palabra</u>**: tales como la longitud, la frecuencia de uso, la regularidad ortográfica, la estructura fonológica o la semejanza con otras palabras, tienen que introducir importantes variaciones en las estrategias de identificación; por ejemplo, es probable que las palabras de elevada frecuencia gocen de una vía más directa de acceso y requieran por ello menos recursos analíticos.

3) **La destreza del lector**: puede facilitar o dificultar la utilización de ciertas estrategias de análisis visual y acceso al léxico; en este sentido, parece obvio que **los lectores expertos utilizan estrategias mucho más automatizadas de identificación de palabras que los lectores novatos** y recurren al análisis de letras en casos excepcionales (vg., para identificar palabras desconocidas o poco frecuentes.
4) **La naturaleza de la tarea**: algunas tareas empleadas en la investigación (vg., la detección de estímulos) resultan extremadamente artificiales y, por eso mismo, pueden requerir el empleo de estrategias ad hoc de las que no se hace uso en la lectura en condiciones naturales; en cambio, otras tareas (vg., el registro de movimientos oculares) no le exigen al sujeto lector la toma de decisiones conscientes que pueden alterar el curso natural de los procesos de lectura.

6.2.- Mecanismos de extracción de la información visual:

El uso de la técnica de registro de los movimientos oculares durante la lectura presenta dos indudables ventajas metodológicas:

a) por un lado, **permite examinar lo que acontece durante los procesos de identificación del lenguaje escrito en unas condiciones comparativamente más naturales**, dado que conlleva el empleo de textos o párrafos, en lugar de palabras aisladas o secuencias de letras sin sentido, lo que hace posible modificar o manipular las distintas variables que a priori pueden intervenir en la actividad lectora en condiciones normales (vg., variables léxicas, sintácticas y semánticas);
b) en segundo lugar, **esta técnica garantiza una medida simultánea o "en curso (on-line) de las actividades cognitivas que acontecen durante la lectura**, con lo que se elude la intervención de variables ajenas a los procesos de lectura que surgen como resultado de decisiones conscientes que se toman a posteriori, es decir, una vez que los procesos responsables de la identificación de unidades lingüísticas han tenido lugar.

Debido a sus ventajas metodológicas, esta técnica de investigación no es privativa de los estudios sobre percepción del lenguaje, sino que se practica también en la investigación de otros niveles de procesamiento lingüístico (vg., comprensión de oraciones y textos).

- Mientras leemos, nuestra mirada se desplaza por el texto en dirección izquierda-derecha, aunque, contrariamente a lo que pudiera parecer, este desplazamiento no es continuo y uniforme, **sino que avanza a pequeños saltos discontinuos, es decir, la mirada no "barre" todas y cada una de las letras del texto, sino que "se posa" en sucesivos puntos del mismo a gran velocidad.** A estas rápidas secuencias de movimientos discontinuos se les denomina movimientos sacádicos.

- Después de cada movimiento sacádico existe un período de duración variable en el que la mirada queda detenida en un punto del texto**; se trata de las llamadas fijaciones oculare**s, que se pueden caracterizar como una especia de instantáneas en las que se registra una región del texto centra en torno al punto de máxima agudeza visual (situado frente a la fóvea) y que abarca un sector periférico ubicado a ambos de dicho punto (el delimitado por la llamada **"visión parafoveal**").

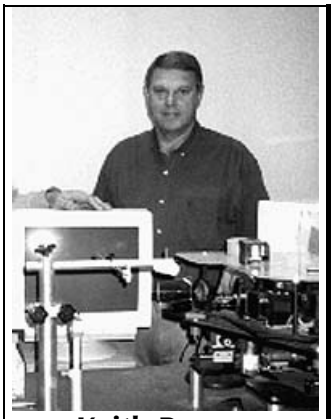
- Los períodos de fijación tienen por objeto extraer la información visual necesaria para la identificación del lenguaje escrito y presentan una duración muy variable **(con un promedio ubicado entre los 200 y los 250 mseg.).**

- Asimismo, **los movimientos sacádicos** presentan una duración variable, dependiendo de la longitud o distancia entre cada dos puntos de fijación, aunque un sacádico típico viene a durar entre **25 y 50 mseg.**

- A la región del texto que queda entre dos puntos de fijación (i.e., a la distancia que cubre cada movimiento sacádico) se le denomina **ámbito de fijación.**

- Además de los movimientos sacádicos, hay otro tipo de movimientos oculares que avanzan en sentido inverso a los sacádicos, esto es, de derecha a izquierda, y que, por tanto, se denominan **regresiones.**

- **Las regresiones** se utilizan, como es obvio para retroceder hacia zonas del texto previamente leídas que requieren una "segunda pasada" o revisión. Estos movimientos regresivos constituyen normalmente entre un 15 y un 20 por ciento del total de movimientos oculares efectuados en la lectura (cfr., Pollatsek y Rayner, 1989, Rayner et al., 1989; Cuetos, 1990).


Keith Rayner

Una primera observación que se desprende de los datos expuestos en el párrafo anterior es que, pese a la ilusión de continuidad y linealidad que experimentamos durante la lectura, **los lectores empleamos cerca del 90 por ciento del tiempo dedicado a leer en detener la mirada en el texto con el objeto de extraer información de él, y tan sólo un 10 por ciento en recorrer el texto en busca de información nueva**. Ahora bien,

¿Cómo extraemos la información visual relevante para identificar el lenguaje escrito?

Como ha señalado Cuetos (1990), hay dos factores que ayudan a esclarecer cómo se extrae la información de la página escrita:

a) Por una parte**, la distribución de fijaciones oculares**, es decir, **los puntos del texto en los que se fija la mirada.** Las variables críticas de este factor son:

- **El ámbito de fijación**, o tamaño de la región que queda entre dos fijaciones sucesivas, y
- **el tipo de unidad lingüística** que es objeto de más frecuentes fijaciones.

b) Por otra parte, el segundo factor **lo constituyen todos aquellos indicios que contribuyen a desvelar los procesos que se ponen en marcha durante cada fijación ocular**, como es el caso de:

- **La duración de las fijaciones** (que se considera un indicio de la dificultad o complejidad de las operaciones cognitivas implicadas en la identificación de unidades lingüísticas) y
- **El ámbito o región del texto que abarca cada fijación ocular** (a través de la visión periférica o parafoveal), que, para distinguirlo del ámbito de fijación, denominaremos alcance perceptivo (perceptual span).

Ámbito de fijación y unidades lingüísticas:

- La amplitud media de los movimientos sacádicos, que determina la región o ámbito de desplazamiento entre ellos, **es de 8 a 10 caracteres** (lo que incluye tanto letras como espacios en blanco), o sea, aproximadamente **una palabra y media**, aunque como puede suponerse, existe una enorme variabilidad de unos a otros. (Rayner et al., 1989).

- El **límite superior del ámbito de fijación** se sitúa en torno **a los 20 caracteres** (Rayner y McConkie, 1976).

- Todo esto hace suponer que **no todas las palabras reciben una fijación ocular.** En general, las palabras cortas (**de hasta 3 letras) son "saltadas" con mayor frecuencia que las largas** (de 6 letras en adelante).

- Asimismo, **las palabras de mayor frecuencia de uso reciben menos fijaciones y de menos duración que las menos frecuentes**.

- Otra distinción relevante a este respecto es la que se da entre palabras funcionales (o de clase cerrada y de contenido (o de clase abierta): aquéllas reciben menos fijaciones y más breves que éstas. Este dato no debe sorprender si se tiene en cuenta que, además de ser en general más cortas, las palabras funcionales suelen ser comparativamente más frecuentes que las de contenido.

Así pues, **en las palabras de clase cerrada o funcionales, la longitud física se alía con la frecuencia y la categoría gramatical para determinar una menor tasa de fijaciones** (del 20 al 40 por ciento); por el contrario, las palabras de contenido o clase abierta, en general menos frecuentes y más largas, presentan una tasa de fijaciones mucho mayor (superior al 80 por ciento) (cfr., Carpenter y Just, 1983; Rayner y Duffy, 1988).

En términos generales, se suele aceptar **la hipótesis de que la probabilidad de efectuar una fijación en una determinada unidad lingüística del texto viene determinada por la relevancia perceptiva o informativa de dicha unidad, o por el grado de dificultad o "profundidad" de procesamiento que requiere su identificación.**

Debido a la naturaleza cuantitativa de esta medida, sin embargo, no resulta fácil desentrañar, en cada caso concreto, cuáles son los motivos específicos que justifican la preferencia por realizar fijaciones en ciertas clases de unidades lingüísticas, ni, por supuesto, calibrar el tiempo de fijación que se destina a cada uno de los diversos subprocesos que integran la comprensión de los mensajes verbales.

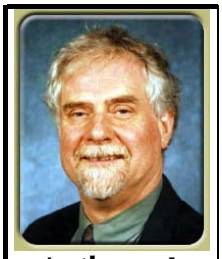
- Así, en el caso del procesamiento visual de palabras, se ha podido constatar la importancia de factores tales como la longitud, la frecuencia de uso o la clase de vocabulario, ya citados, y de otros distintos, como la predictibilidad o la ambigüedad de la palabra.

- Asimismo, se ha descubierto que los verbos principales de las oraciones suelen registrar más fijaciones y de mayor duración, y que lo mismo sucede con las palabras situadas al término de unidades estructurales y semánticas (vg., la cláusula o la oración).
- Por ejemplo, en un reciente estudio de Rayner et al. (1989), se comprobó que **la duración de la fijación en una determinada palabra (vg., la palabra "planeta") era sensiblemente mayor cuando dicha palabra se hallaba al final de una cláusula subordinada (ejemplo 2-a), que cuando se encontraba en mitad de la cláusula (ejemplo 2-b), aun cuando las dos oraciones fueran idénticas en todo lo demás.**

2.-

a) Mucho antes de que aparecieran los primeros homínidos en nuestro planeta, ya había formas más elementales de vida.
b) Mucho antes de que aparecieran los primeros homínidos, en nuestro planeta ya había formas más elementales de vida.

- Este efecto se ha interpretado como una manifestación de los procesos de integración de información que tienen lugar al descodificar unidades lingüísticas con consistencia sintáctica y semántica (vg., cláusulas y/u oraciones), por lo que no se trata en absoluto de un efecto léxico, sino de un fenómeno oracional que aflora, de forma más bien accidental, en el plano de las palabras.



Anthony J. Sandford

Fenómenos similares a éste se han observado en **estudios de registro de movimiento oculares relativos a la búsqueda de antecedentes de expresiones anafóricas** (cfr., Sanford y Garrod, 1981) **y a la realización de inferencias elaborativas durante la comprensión de textos escritos.** (cfr. O'brien et al., 1988).

Alcance perceptivo e integración de información en fijaciones sucesivas:

Uno de los inconvenientes más señalados de la técnica de registro de movimientos oculares es, sin duda, **su imprecisión en lo que concierte a las causas de la duración de las fijaciones.** Bien es verdad, no obstante, que este defecto perjudica por igual a todos los procedimientos de recogida sistemática de datos que se basan en índices de tiempo, de los que tenemos una buena cantidad en psicología del lenguaje, (vg., medidas de latencia en la comprensión, pausas en la producción, etc.).

Ya se ha visto que una parte aparentemente sustancial del tiempo de fijación se consume en procesos de niveles superiores (vg., léxicos, sintácticos y semánticos).

Algunos autores (cfr. Rayner et al., 1981) **cifran en 50 mseg.** el lapso de tiempo mínimo necesario para extraer información visual en la lectura, lo que representa entre la cuarta y la quinta parte del tiempo total de cada fijación. Ahora bien, con independencia de estas precisiones, interesa determinar qué es lo que ocurre durante el tiempo de fijación.

- Un aspecto a destacar de los movimientos sacádicos es que **son de naturaleza "basilística",** lo que supone que, una vez iniciados, no admiten correcciones sobre la marcha. Por ello, el punto de destino de cada movimiento sacádico debe estar establecido de antemano antes de que éste se inicie.

- La única manera de explicar esta especie de "don adivinatorio" de los movimientos sacádicos es suponer, como han ahecho algunos autores (Rayner, 1975; McConkie y Rayner, 1975), **que el sistema visual tiene acceso a la información situada por delante (o a la derecha) del punto de fijación y utiliza esta información para programar el alcance del siguiente desplazamiento.**

- El alcance perceptivo medio de la visión periférica en cada fijación es de **unos 4 ó 5 caracteres a la izquierda** del punto de fijación y de **14 caracteres a su derecha,** lo que supone que, en condiciones normales, el alcance perceptivo de un lector se extiende más allá de la palabra situada en su "punto de mira".

- En efecto, se ha comprobado (cfr. Rayner et al., 1982) que el ritmo de lectura es más veloz cuando al lector se le muestran simultáneamente la palabra que ocupa el punto de fijación y la que le sigue, que cuando tan sólo se le muestra en pantalla la palabra fijada.

La técnica empleada para desvelar el tamaño de la región que abarca cada fijación es la llamada **"técnica de la ventana móvil"** (McConkie y Rayner, 1975; Pollatsek y Rayner, 1989), una técnica que, como se verá, ha sido empleada también en otros menesteres. Brevemente, esta técnica consiste en acotar una región de texto alrededor del punto de fijación y mantenerla constante a medida que el sujeto va desplazando su mirada a lo largo del texto.

- La amplitud de esta "ventana móvil" y su localización durante los movimientos sacádicos son establecidas y controladas por el experimentador. Así, dentro de esta ventana definida por el experimentador, el sujeto dispone sólo de un fragmento del texto; la parte del mismo que queda fuera de la ventana es mutilada o alterada de diferentes maneras.

- Un dato interesante recogido mediante el uso de la ventana móvil es que cuando la región exterior a dicha ventana se halla ocupada por letras y espacios asignados aleatoriamente (ejemplo 3-a), **los sujetos rara vez son conscientes de los cambios en el desplazamiento a lo largo del texto o del contenido de esta región exterior,** y creen hallarse ante un texto normal; en cambio, si la región exterior está formada por una secuencia de X (ejemplo 3-b), los lectores sí distinguen entre la porción de texto interna a la ventana y la zona exterior a la misma (el asterisco indica el punto de fijación; la oración completa sería: "Con mucho, la sustancia simple más abundante en la biosfera...").

3.-

a) Oqm kxemw, fa sustancia simple zif ewjpsldyf vr pq copzmwlh.
b) XXX XXXXX, Xa sustancia simple XXX XXXXXXXXX XX XX XXXXXXXX

- Para Pollatsek y Rayner (1989), esto revela que el alcance perceptivo puede exceder la región percibida por la visión periférica. Puestas así las cosas, es necesario preguntarse qué clase de procesamiento es el que se aplica a la información registrada a través de la visión parafoveal y cómo se conjuga esta información con la que se obtiene en la fijación siguiente; dicho en otras palabras,

¿Cómo se integra la información obtenida en dos fijaciones sucesivas?

Rayner (1975) intentó responder a esta pregunta empleando un ingenioso procedimiento denominado **"técnica de los límites (boundary technique)**.

- El procedimiento consiste en seleccionar una palabra crítica de un texto que ha sido objeto de una fijación durante la lectura (vg., "boda") y reemplazar esa palabra por otra (vg., "cera") o por cualquier otro estímulo gráfico (vg., una no-palabra o una secuencia impronunciable de letras).

- El estímulo sustituyente ("cera") se presenta inicialmente (i.e., en el momento de la primera fijación o fijación n) para ser reemplazado por la palabra original ("boda") en la segunda fijación (o fijación n + 1).

- De esta manera, el sujeto **no percibe conscientemente ninguna alteración en el texto**; sin embargo, si el estímulo "inductor" ha sido procesado en la visión periférica deberá producir efectos sobre la duración en la fijación n +1 (en la que la palabra original aparece en la misma posición del texto). El ejemplo 4 ilustra las dos fijaciones junto con los puntos de fijación (1 y 2) y un hipotético límite entre las dos fijaciones (indicado por una barra vertical):

4.-
- Fijación n: El camarero llevó corriendo el pastel de cera a la mesa.
1 | 2
- Fijación n+1: El camarero llevó corriendo el pastel de boda a la mesa

- Los experimentos en los que se empleó la técnica de los límites (Pollatsek y Rayner, 1989) **ofrecen resultados contradictorios**, aunque, en líneas generales, parecen indicar que cuando el estímulo sustituyente comparte con la palabra original las primeras letras (vg. boxx y boda, o bota y boda), se da un efecto de facilitación, es decir, una reducción en el tiempo de fijación de la palabra crítica, efecto que se mantiene aun cuando la caja sea distinta en ambos estímulos (vg., BODA-boda, o bOdA-BoDa);

Esto indica que **la visión parafoveal permite un procesamiento relativamente abstracto de los estímulos visuales.** No obstante, cuando la semejanza entre ambos estímulos es semántica (vg., coche-auto, pelea-lucha) **el efecto facilitador desaparece por completo**.

En suma, como concluyen Pollatsek y Rayner, "no hay pruebas de que exista significación funcional en la codificación parafoveal de otros estímulos que no sean las representaciones abstractas de las letras"

6.3.- La percepción del lenguaje oral y escrito: algunas conclusiones.

A través del repaso sumario a las fases iniciales del reconocimiento del lenguaje escrito ofrecido en este apartado, se pueden destacar unos pocos puntos de convergencia entre los procesos de identificación del habla y del lenguaje escrito, y algunos puntos más de divergencia.

a) Entre los primeros, hay que **señalar la presencia de niveles de procesamiento similares en ambas modalidades.**

- Los procesos de percepción del lenguaje tienen un cometido común, a saber, **la obtención de una representación de los elementos léxicos que integran el mensaje verbal.**
- Este camino hacia el léxico, apenas iniciado aquí, atraviesa una serie de etapas intermedias caracterizadas por **la recuperación de unidades lingüísticas subléxicas (que van desde los rasgos acústicos o visuales, en un extremo hasta las representaciones fonológicas u ortográficas en el otro).**
- Aquí terminan prácticamente las semejanzas entre el reconocimiento auditivo y visual del lenguaje.

b) Entre las divergencias más notorias, es preciso reconocer **el distinto papel que se debe atribuir a las características físicas** (i.e., las claves acústicas y los rasgos visuales de la entrada sensorial, que en el caso del lenguaje escrito, no parecen suscitar dificultades de decodificación en integración comparables a las que plantea el lenguaje oral.

- **Problemas como el de la ausencia de invarianza o la segmentación fonética no parecen existir en el dominio del procesamiento visual.**

Bien es verdad, no obstante, que nos hemos limitado a examinar aquí la modalidad de lenguaje escrito que menos problemas de segmentación e identificación origina, el lenguaje impreso. Parece fuera de toda duda **que un estímulo visual como la escritura manual, que por su propia naturaleza es mucho más variable e irregular que el lenguaje impreso, tiene que plantear forzosamente problemas de segmentación muy similares a los que suscita el lenguaje oral.**

En este sentido, lo mismo que la percepción de sonidos del habla es un proceso íntimamente dependiente de los efectos de coarticulación, así también podemos suponer **que la percepción de la escritura manual tiene que depender de manera crucial, y en mayor medida que la percepción de la escritura impresa, de factores contextuales.**
Como consecuencia de ello, podríamos incluso anticipar la hipótesis de que el efecto de superioridad de la palabra sea más consistente en la percepción de la escritura manual que en la de la escritura impresa.

Una segunda divergencia entre las dos modalidades lingüísticas de entrada (visual y auditiva) radica **en la mayor relación de interdependencia que hay entre la percepción y la producción del habla, frente a una más clara autonomía funcional de los procesos de lectura con respecto a la escritura, excepción hecha, naturalmente, de la percepción de la escritura manual, por las razones que acabamos de apuntar.**

En este último apartado del capítulo no se ha hecho referencia alguna a los procesos de escritura. La omisión ha sido intencionada, ya que la escritura es una habilidad relativamente tardía en su desarrollo (inclusive con respecto a la lectura) y, sobre todo, se halla considerablemente encapsulada, al menos en sus aspectos más específicos, es decir, en el plano del control motor. Esto ha justificado tradicionalmente su apartamiento de los textos de psicología del lenguaje y su consideración relativamente independiente

entre los temas de esta disciplina.

La cuestión de las relaciones entre las modalidades visual y auditiva de procesamiento del lenguaje, sin embargo, no se agota aquí. El próximo capítulo será escenario de uno de los problemas recurrentes de la psicolingüística moderna, el que se refiere a las rutas de acceso léxico mental en el lenguaje oral y en el lenguaje escrito. Dejando, pues, el camino hacia niveles superiores de procesamiento en la comprensión del lenguaje.